A man would do nothing, if he waited until he could do it so well that no one would find fault with what he has done.

CARDINAL NEWMAN

University of Alberta

Studies on Surface-bound Molecular Architectures as Components of Nanodevices

By

Sulayman Adeyemi Oladepo ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

Department of Chemistry

Edmonton, Alberta

Fall 2006

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-22339-0
Our file Notre référence
ISBN: 978-0-494-22339-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The fabrication and characterization of molecular architectures is now a major area of interest of contemporary chemistry. There is however the challenge of developing new methods for their fabrication in a rapid, facile and reproducible manner. The research work presented in this thesis explored the fabrication and spectroscopic characterization of single-stranded deoxyribonucleic acid (ssDNA) on one hand and computational modeling of the mechanical and electronic properties of polyyne molecules on the other hand. Both of these molecular units are building blocks which can be used to create functional constructs of desirable properties and applications.

We immobilized ssDNA molecules on gold surface. These were then characterized with X-ray photoelectron spectroscopy (XPS) and Fourier-transform infrared spectroscopy (FTIR). Both techniques gave some evidence on the adsorption of the DNA molecules on the gold surface, as well as the nature of the adsorption in terms of specific or non-specific chemisorption.

In a similar and related effort to explore properties of molecular architectures, we used computational approach to model the mechanical and electronic properties of polyyne molecules. These molecules are believed to possess unique electronic and optical properties; hence, they may be used as components of molecular electronic devices. Computations at different levels of theory, from semi-empirical PM3 to DFT were used to gain insight into the mechanical and electronic properties of these molecules. The results obtained and the inferences made therefrom are presented.

Acknowledgements

In the first place, I give thanks to Almighty Allah for seeing me this far in my academic pursuit. I invoke His blessings on the gentle soul of the seal of prophets, Muhammad (SAW).

I wish to express my heartfelt gratitude and appreciation to my M. Sc. Supervisor, Dr. John-Bruce D. Green for his thorough supervision, support and guidance. I must say without mincing words that I enjoyed your painstaking criticisms, comments and above all, your overall friendly style of supervision. The memory of the time spent with you will tarry with me for a long time.

It will be a great ingratitude if I fail to acknowledge the support, love, understanding and perseverance of a dear and unique friend, who, by marital protocols is my wife, Bilqees Abimbola Oladepo. Precious, I will forever be grateful for finding a friend in you and for being there for me at all times!

At this point, I wish to thank my dear parents, Mallam Imran Alao Oladepo and Madam Nimotallahi Abike Oladepo, for their support and understanding. I pray Allah to preserve both of you in good health so that you can reap the fruits of your hard labour. I also thank my siblings, Mr. Musodiq Oladepo and family, Mr. Qazeem Oladepo and family, Mrs. Monsurat Hammed and family, Basirat Oladepo and Sakirat Oladepo for their unflinching support and understanding. I must also thank my very good friend, Mr. 'Wale Yussuff and all other family members, friends and associates in Nigeria for their support and interest, May Allah reward all of you in manifolds.

I express my appreciation to Dr. Gino DiLabio of the National Institute for Nanotechnology (NINT) for helpful discussions and Dr. Alex Brown of Chemistry department for permission to use computing facility. I am grateful to the Faculty of Graduate Studies and Research, University of Alberta, as well as the Natural Sciences and Engineering Research Council of Canada (NSERC), for providing the funding for my research projects. I also thank members of staff of Chemistry department for their support, and the university community as a whole for providing

such an enabling environment for my academic pursuits. I also thank in a special way, Mr. Francis Nsiah, for his avuncular advice and support. Thank you for being a wonderful friend and brother!

It is but pertinent to recognize the contributions of my M. Sc. Committee members. I appreciate your comments and criticisms. I also thank all members of John-Bruce Green's group, both past and present, for their various contributions. In addition, I want to thank my sisters and friends in Edmonton, Khadijah and Toks, as well as all my friends, colleagues and associates in Canada, the US and UK, for their interest and support. I am indeed very grateful to you all.

Sulayman Oladepo
May 2006.

Dedication

Dedicated to my wife Bilqees Abimbola Oladepo…
and to the memory of my late granny, Madam Asmau Oladepo (RIP) for her support in her days.

Table of Contents

CHAPTER 1: Introduction

PAGE

Molecular Architectures 1
Fabrication of Molecular Architectures 2
Self-Assembled Monolayers (SAMs) 3
Characterization of SAMs 6
Immobilization and Characterization of ssDNA on gold 7
Computations 11
Research Objectives 13
References 15

CHAPTER 2: Spectroscopic Characterization of Oligonucleotides Immobilized on Gold

Introduction 19
Experimental 21
Results and Discussion 24
Conclusions 42
References 43

CHAPTER 3: Exploration of the Performance of Selected DFT Functionals for Predicting Vibrational Frequencies of 1, 12-Diphenyl-1, 3, 5, 7, 9, 11-dodecahexayne

Motivation 45
Introduction 45
Calculation Method 48

Results and Discussion 51
Conclusions 72
References 74

CHAPTER 4: Modeling the Mechanical and Electronic Properties of Polyyne Molecules: A Density Functional Theory Approach

Introduction 77
Calculation Method 80
Results and Discussion 86
Conclusions 116
References 118

CHAPTER 5: General Conclusions 121
References 127

APPENDIX 129
Computation Methods 129
DFT Functionals Used for Calculations 130
Vibrational Frequencies 132

List of Tables

Table 2.1: The base sequence of the two main ssDNA samples 23

Table 2.2: A brief description of the samples used 23

Table 2.3: Peak areas (in millions) of the major XPS signals of the five samples including the control 25

Table 2.4: Results of the film parameters t and N_N/ N_{Au} as obtained using XPS results and the SRD-82 software 38

Table 3.1: The results obtained from previous calculations as compared with reported experimentally observed values 53

Table 3.2: The results obtained from our calculations with **B3LYP/6-31G***) as compared to reported experimental and theoretical literature values 55

Table 3.3: The results obtained from our calculations with **BLYP/6-31G*** as compared to reported experimental and theoretical literature values 56

Table 3.4: The results obtained from our calculations with **BPW91/6-31G*** as compared to reported experimental and theoretical literature values 57

Table 3.5: The results obtained from our calculations with **B3LYP /6-31G**** as compared to reported experimental and theoretical literature values 59

Table 3.6: The results obtained from our calculations with **BLYP/6-31G**** as compared to reported experimental and theoretical literature values 60

Table 3.7: The results obtained from our calculations with **BPW91/6-31G**** as compared to reported experimental and theoretical literature values 61

Table 3.8: The overall deviations of our values from experimentally observed values as obtained for each of the methods 62

Table 3.9: Results obtained with the fifteen DFT functionals on the model compound (DP-C12) 65

Table 3.10: The results obtained from our calculations with **BLYP/6-31G****as compared with reported experimental values 68

Table 3.11: The estimated effective bending and stretching force constants for C_4-C_{12} polyyne molecules 71

Table 4.1: The force constants obtained from the bending of selected polyynes 93

Table 4.2: The stretching force constants obtained for C_4-C_{40} polyynes 97

Table 4.3: The bending force constants obtained for the selected members of the polyyne series from DFT calculations 106

Table 4.4: The stretching force constants obtained for the selected members of the polyyne series from DFT calculations 109

Table 4.5: HOMO-LUMO Egap for the five polyyne molecules 115

List of Figures

Figure 1.1:	Schematic of an alkanethiol adsorbed onto a gold substrate via the thiol head group, with the alkyl tail group pointing outwards from the gold surface	3
Figure 1.2:	Schematic of an ssDNA immobilized on a gold substrate via the thiol group	9
Figure 2.1:	High-resolution XPS spectra of O1s peak region for all the five samples	30
Figure 2.2:	High-resolution XPS spectra of C1s peak region for all five samples	31
Figure 2.3:	High-resolution XPS spectra of N1s peak region for the five samples	32
Figure 2.4:	High-resolution XPS spectra of P2p peak region for the five samples	33
Figure 2.5:	High-resolution XPS spectra of S2p peak region for all five samples	34
Figure 2.6:	FTIR spectrum of 1μM TT in 1TE-NaCl buffer	40
Figure 2.7:	FTIR spectrum of 1μM AA in 1TE-NaCl buffer	41
Figure 3.1:	Cartoon showing the structure of ethylene molecule and some possible vibrational modes	52

Figure 3.2:	The chemical structure of the model molecule DP-C12	64
Figure 3.3:	A typical polyyne molecule shown as being made up of constituent units	70
Figure 4.1:	Model of a polyyne molecule bent through 30°	82
Figure 4.2:	A cartoon showing the stretching simulation	84
Figure 4.3:	Plots of offset total energy vs bending angles for C_4-C_{20}	86
Figure 4.4:	Plots of offset total energy vs equivalent length for C_4-C_{20}	87
Figure 4.5:	Plots of C-C-C bond angles for C_4-C_{20} bent at 20°	89
Figure 4.6:	Plots of force vs equivalent length for C_4-C_{20}	92
Figure 4.7:	Plot of force constants vs number of carbon atoms per molecule of polyyne	94
Figure 4.8:	Plots of offset total energy vs stretching coordinate for C_4-C_{40}	95
Figure 4.9:	Plot of force constant vs number of carbon atoms per molecule for C_4-C_{40}	98
Figure 4.10:	Plots of offset force vs % stretched molecular length ($\%\nabla L / L_o$) for C_4-C_{40}	99

Figure 4.11: Plot of the slope of offset force against % stretched molecular length vs number of carbon atoms per molecule for C_4-C_{40} 100

Figure 4.12: Plots of offset total energy vs equivalent length for C_4-C_{12} polyynes 102

Figure 4.13: A cartoon of 1, 3, 5, 7-octatetrayne molecules bent at 20° and geometry optimized by PM3 and DFT 103

Figure 4.14: This figure shows the plots of C-C-C bond angles for C_4-C_{12} bent at 20° 105

Figure 4.15: Plot of force constant vs number of carbon atoms per molecule for C_4-C_{12} polyynes 107

Figure 4.16: Plots of offset total energy vs stretching coordinate for C_4-C_{12} 108

Figure 4.17: Plot of stretching force constant vs number of carbon atoms per molecule for C_4-C_{12} 110

Figure 4.18: Plots of offset force vs % stretched molecular length $(\%\nabla L / L_o)$ for C_4-C_{12} 111

Figure 4.19: The HOMO-LUMO Egap normalized to the Egap at 0° bending angle 112

Figure 4.20: Normalized Egap plotted as a function of average atomic angle 114

List of Symbols

Å	Angstrom (a unit of length equivalent to 10^{-10} m)
β	Asymmetry parameter for electrons
c	Velocity of light in cm/s
Δ	Relative phase change
∇L	Change in length
E	Young's modulus
E_g	Band gap energy
E_l	Equivalent length
E_v	Experimentally observed wavenumber
I	Moment of inertia
I_{Au}	Signal intensity for Au
$I^o{}_{Au}$	Intensity (peak area) of a clean gold substrate
I_N	Signal intensity for N
θ	Bending angle in degrees
θ_a	Advancing contact angle

θ_N	Number of N atoms per unit area of film
θ_N/N_N	Relative N coverage
k	Force constant in N/m
K_{eff}	Effective force constant
k_n	Force constant of a unit spring
L	A measure of the length of a rod or a polyyne molecule
L_{Au}	Average practical EAL (PEAL) for electrons from gold
L_N	Average practical EAL (PEAL) for electrons from DNA in the DNA overlayer
L^Q_{Au}	Effective attenuation length for quantitative analysis for Au
L^Q_N	Effective attenuation length for quantitative analysis for N
Lo	Original unstretched length
L_l	Stretched length
μ	Reduced mass
n	Number of carbon atoms per molecule
n_d	Number of $C \equiv C$ units joined together in series

n_{DNA}	Absolute coverage of DNA
N_N	Atomic density for N
N_N / N_{Au}	Relative atomic density for nitrogen
ν	Wavenumber in cm^{-1}
O_l	Original length
S_l	Stretched length
σ_{Au}	Total photoelectric cross section for Au
σ_N	Total photoelectric cross section for N
Ψ	Relative amplitude change
t	Film thickness to be determined
T_{Au}	Analyzer transmission function for Au
T_N	Analyzer transmission function for N
$V_{c,}$	Calculated wavenumber
$\#_{vib}$	Number of vibrational modes
$\#$	Number

%	Percentage

List of Abbreviations

AA	A sample of non-thiolated ssDNA containing adenine of 20 bases
AFM	Atomic force microscopy/microscope
AICT	Academic Information and Communication Technology
AA/MCH	AA backfilled with MCH
AM1	Austin Model 1
AO	Atomic orbital
Asym-str	Asymmetric stretch
a.u.	Absorbance units
BE	Binding energy
C	Carbon
c-AFM	Conductive atomic force microscopy
c.eV	counts-electronvolt
CGTF	Contracted Gaussian-type functions
C-H	Carbon-hydrogen bond

C-N	Carbon-nitrogen bond
CNS	Computing and Network Services at the University of Alberta
Cr/Au	Chromium/gold
^{13}C NMR	Carbon-13 nuclear magnetic resonance
DDE	Dynamic Data Exchange
DFT	Density functional theory
DNA	Deoxyribonucleic acid (oligonucleotides)
DP-C12	1, 12-Diphenyl-1, 3, 5, 7, 9, 11-dodecahexayne
DZ	Double-zeta
DZP	Double-zeta plus polarization
EAL	Effective attenuation length
EDTA	Ethylenediamminetetraacetic acid
Egap	HOMO-LUMO Energy gap
eV	Electronvolt
FTIR	Fourier-transform infrared spectroscopy
GGA	Generalized-gradient approximation

GTFs	Gaussian-type functions
H	Hydrogen
He	Helium
HF	Hartree-Fock
HOMO	Highest occupied molecular orbital
IR	Infra-red
kcal/molÅ^2	Kilocalorie per mole angstrom squared
KS	Kohn-Sham
LB	Langmuir-Blodgett
LCAO-SCF	Linear combination of atomic orbitals – self-consistent field
LDA	Local-density approximation
LSDA	Local-spin-density approximation
LUMO	Lowest unoccupied molecular orbital
MCH	6-mercapto-1-hexanol
Mg and Al	Magnesium and Aluminium
mM	Millimolar

MPAD	Mean percentage absolute deviation
MP2	Second-order Møller-Plesset perturbation theory
μL	Microlitre
μm	Micrometer
μM	Micromolar
N	Nitrogen
NIST	National institute of standards and technology
N/m	Newton per meter
O	Oxygen
P	Phosphorus
PAD	Percentage absolute deviation
PBS	Portable batch system
PEAL	Average practical EAL
PM3	Parametric method 1
QEAL	Quantitative EAL
QC-SCF	Quadratically convergent self-consistent field

-R	Alkyl tail group of an alkanethiol
RHF	Restricted Hartree-Fock
RPA	Random phase approximation
S	Sulphur
SA	Self-assembly
SAM	Self-assembly monolayers
SCF	Self-consistent field
Scis	Scissoring
Sc-Zn	From Scandium to Zinc
S-H	Thiol-hydrogen bond
Si	Silicon
SPM	Scanning probe microscopy
SPR	Surface plasmon resonance
SQM	Scaled quantum mechanical
SRD-82	Standard reference database 82
ssDNA	Single-stranded DNA

STM	Scanning tunneling microscopy
STOs	Slater-type orbitals
Str	Stretch
Sym-str	Symmetric stretch
TT	A sample of thiolated ssDNA containing thymine of 20 bases
TE	Tris-EDTA
TT/MCH	TT backfilled with MCH
ToF-SIMS	Time-of-flight secondary ion mass spectrometry
UV	Ultraviolet
XPS	X-ray photoelectron spectroscopy
# *Natoms*	Number of N atoms
$(\%\nabla L / L_o)$	Percentage stretched molecular length

CHAPTER 1

Introduction

Molecular Architectures

The realization that simple molecular units can serve as building blocks that contain the information needed to assemble an astounding array of new chemical architectures has captured the imagination of chemists across the discipline [1]. Therefore, the manipulation of weak forces to construct new molecular architectures is now a major focus of contemporary chemistry [2]. The term "molecular architecture" seems ambiguous [3], however, it may be defined as an orderly arrangement of functional molecular units, which serve as building blocks to create a new supramolecular entity with unique properties and applications. Thus, simple and beautiful molecular structures can be converted into functional constructs (molecular architectures) [1]. The building blocks might include many simple and familiar molecules [1] such as porphyrins, fullerenes, alkanethiols, biomolecules, multifunctional heterocycles, metal complexes etc. It has become apparent that the macroscopic properties of such architectures are directly related to their microscopic structure [4]. Thus, the properties of the building blocks can be favourably exploited in order to produce a functional construct with desirable properties and applications. The potential applications of these constructs include use in microarrays, chemical and biochemical sensing, gene mapping, molecular recognition, medical diagnostics, molecular electronics, photonics, non-linear optics, materials science, nanoscience and nanotechnology [1, 5-14]. With these potential applications in mind, the construction of thin film architectures with novel properties has been the subject of intense exploration, and the challenge is to develop new methods for their construction in a rapid, facile and reproducible manner [4].

Substrate-supported molecular architectures of polyynes have not been successfully fabricated though; their properties are worth exploring due to their unique electronic and optical properties [10, 15-22]. These properties can be appropriately exploited for different applications such as molecular electronics [10, 23], nonlinear optics [10], nanotechnology [13], biology and material science [19]. This explains why these compounds have been the subjects of much research interest over the years [10, 12, 16-20]. These research efforts are aimed at exploring the possible properties of these compounds as well as their potential applications. Much of the work on potential applications is computational. Since the bulk properties of an array of these compounds are directly related to the microscopic properties [4], most of the computation work is carried out on single molecules [10, 12, 24]. Rational deductions from the results of these computations will aid our understanding of the unique properties of these compounds, the way in which their molecular architectures can be fabricated and the range of applications to which they may be put.

Fabrication of Molecular Architectures

Molecular architectures can be constructed in a variety of ways, depending on properties of the building blocks and the desired applications. However, two common methods are the Langmuir-Blodgett (LB) techniques and self-assembly (SA) [14]. Although LB films can yield densely packed, highly ordered monolayers (single molecular layers) [25] and multilayer films, they require specialized equipment, are limited to substances derived from amphiphilic molecules, and are often easily removed from the substrate [4]. SA is an attractive alternative to LB because uniform monolayers can be spontaneously chemisorbed onto a substrate by immersion in dilute solutions of an adsorbate [4]. Thus, because of this attractive feature, the process of self-assembly has been embraced by many and is being examined on a wide variety of systems such as thiols and disulfides on gold [26-29]. Of note is the fact that the self-assembled monolayers (SAMs) of alkanethiols on gold have been extensively studied [5, 25]. The mechanisms and kinetics of SAMs formation in general have also been a subject of excellent review [25]. Thus, in the next section, we focus on the process of self-assembly and how it is used to form monolayers.

Self-Assembled Monolayers (SAMs)

As stated in the preceding section, SAMs are now widely formed by immersion of substrates in dilute solutions of the desired adsorbates. Here, our focus will be on SAMs of functionalized thiol-derivatized molecules on gold, using alkanethiols as representative example. Thiol-derivatized molecules have very strong affinity for transition metal surfaces, but the most studied and perhaps the most understood SAM is that of alkanethiol on gold surfaces [14]. In principle, a molecule is given a head group, which has strong affinity for the substrate surface (e.g. thiol head group and gold surface). The thiol-derivatized molecules adsorb onto the gold from a dilute solution forming a dense monolayer with the tail group pointing outwards from the gold surface (see Figure 1.1). The surface functionality can be greatly varied by changing the functionalities of the tail groups as desired. There is also the possibility of functionalizing the tail groups by performing reactions on the tail groups after forming the SAMs. Due to the specific nature of the molecule-substrate interaction, the adsorption process can be carried out in a variety of solvents – polar and non-polar.

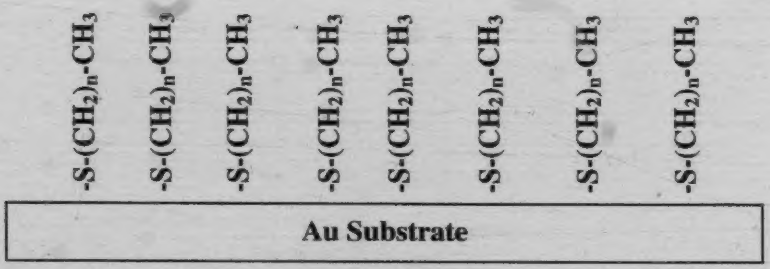


Figure 1.1: Schematic of an alkanethiol adsorbed onto a gold substrate via the thiol head group, with the alkyl tail group pointing outwards from the gold surface. The subscript n can be any whole number depending on the desired chain length.

The rate of formation of SAMs is influenced by several factors such as temperature, solvent, concentration and chain length of the adsorbate, and cleanliness of the substrate [28]. Kinetic studies on adsorption of alkanethiols on gold surface indicate that, at moderate concentrations (10^{-3}M), the adsorption process is composed of two stages [14, 28, 30]. A very fast step (solution-phase bulk transport and adsorption of molecules) [15], which takes a few minutes, by the end of which the contact angles (used in contact angle measurements) are close to their limiting values and the thickness about 80-90% of its maximum value (thickness attained at the end of adsorption), and a slow step, which lasts several hours, at the end of which the thickness and contact angles reach their final values. The initial fast step which is described by diffusion-controlled Langmuir adsorption is strongly dependent on thiol concentration [14, 28], while the second slow step can be described as a surface crystallization process, where the tail groups get out of the disordered state and into unit cells, thereby forming a two-dimensional crystal (self-organization on the surface) [14, 25]. Thus, the kinetics of the first step is governed by surface-head group interaction, while that of the second step is related to chain disorder, chain-chain interactions and the surface mobility of chains [14]. In addition, it was found that the kinetics is faster for longer alkyl chains than for shorter chains [14, 28]. In contact angle measurements, the parameter used is the contact angle, which is a quantitative measure of the relative hydrophobicities of surfaces when wetted by a liquid droplet. Since it is believed that the liquid directly comes in contact with the tail groups of the alkanethiols, this method can give information on the nature of the monolayer on the surface. The higher the contact angle, the more hydrophobic the surface is. The advancing contact angle of water and hexadecane [$\theta_a(H_2O)$ and $\theta_a(HD)$] have been used as criteria for the formation of a good monolyer by some researchers [20]. Long-chain, methyl-terminated molecules are considered by these authors to form a well-packed monolayer when the $\theta_a (H_2O) > 100^\circ$ and $\theta_a (HD) > 40^\circ$ [20].

The chemisorption of alkanethiols on gold surfaces yields gold(I) thiolated (RS^-) species [14, 30]. Though the mechanism involved in the reaction of alkanethiols with gold is not completely understood [30], the reaction may be considered as an oxidative addition of the S-H bond to the gold surface, followed by a reductive elimination of the hydrogen. The proton probably ends up as H_2 molecule, as given by the following reaction [14, 30]:

$$\mathbf{R\text{-}S\text{-}H + Au^0_n \longrightarrow R\text{-}S^-\text{-}Au^+ + 1/2H_2 + Au^0_n}$$

One fascinating feature of SAMs formation is the fact that mixed monolayers can be formed. This may be achieved by mixing the adsorbates of different functionalities in solution, so that after the self-assembly process, the surface will be composed of two functionalities. In this case, the relative proportion of the different functionalities on the gold surface will depend on such factors as mixing ratio of adsorbates in solution, solubilities of the thiols in the chosen solvent and the nature of the tail groups. Another way in which mixed monolayers may be achieved is by backfilling the already formed SAM with a solution of another thiol such that the new thiol can displace some of the already adsorbed molecules. This method has been used extensively in the immobilization of oligonucleotides on gold [7], where the surface with immobilized oligo is immersed in a solution of 6-mercapto-1-hexanol (MCH) resulting in mixed monolayers of the oligo and MCH.

Another related and similar approach to the formation of SAM is by vapour deposition. One way of doing this is that reported by Chailapakul [29], where vapour-deposited SAMs were prepared by exposing the gold surface to n-alkanethiol vapours followed by nitrogen purging, which removes loosely bound molecules from the surface. Experimental results have shown that the SAMs formed from both liquid and vapour phases are identical in structure. It has also been shown that the thickness, structure and packing densities of SAMs deposited from the two phases are indistinguishable [29]. By packing density, we mean how closely packed the molecules are on the surface. However, higher resolution scanning tunneling microscopy (STM) revealed some differences between the two SAMs. For instance, it was observed that thiol-induced changes in gold morphology are different depending on the contacting ambient phase. In STM, an atomically sharp metallic tip is brought in close proximity with a conductive surface, with an applied voltage between the tip and the surface. A tunneling current flows when the tip-surface distance is small enough, and this current is exponentially related to the tip-sample separation. There will be an exponential change in the tunneling current as the tip scans the surface and encounters features of varying heights. The current is kept constant by the vertical movement of the piezoelectric scanner whose changing heights are converted into voltage to give the topographic image of the surface.

Characterization of SAMs

After successful formation of SAMs with desired functionalities, the monolayer can be characterized by a host of surface analytical tools. This is essential in order to ascertain that the structure and properties of the SAMs so created meet the desired qualities. Some of the most frequently used techniques are electrochemistry, scanning probe microscopy (SPM), ellipsometry, contact angles, Fourier Transform Infrared Spectroscopy (FTIR) and X-ray photoelectron spectroscopy (XPS) [14, 25, 28-33].

Electrochemical characterization of SAMs could be in different forms, depending on the interest of different researchers. For instance, J-B. D. Green and coworkers have used AFM-based adhesion measurements to monitor in real time the electrochemical transformation of a surface-bound redox species. They obtained force curves during the cyclic voltammetric scans and then correlated the changes in the tip-sample force of adhesion with the redox transformation of the redox species on the gold electrode [31]. Other researchers have used this method to probe phenomena like electrochemical nanostructuring of Au surfaces, relationship between electrode activity and surface topography, as well as investigating the controlled deposition of metals from solutions by way of deposition morphology and mechanism [32].

SPM is a family of related techniques [e.g. atomic force microscopy (AFM) and scanning tunneling microscopy (STM)], all of which are based upon scanning a probe (tip) above a given surface while monitoring the tip-sample interactions. The techniques are capable of providing information on the surface structure at the atomic scale. In AFM, the tip scans the surface in a raster pattern and the feedback mechanism causes the piezoelectric scanner to maintain the tip at a constant force or height. With the optical detection system, a laser is focused on the back of the cantilever and as the tip scans the surface, the laser beam is reflected off the cantilever onto a position-sensitive phodetector. This detector measures the light intensities (based on the up and down movements of the tip on the surface), converts these into voltage and generate a topographic image of the surface.

Ellipsometry is another surface-sensitive technique in common use. It is based on the measurements of the changes in the polarization of light incident on a surface at an angle. The relative phase change Δ and relative amplitude change Ψ are determined, from which film parameters such as film thickness and morphology can then be deduced. Detailed description of FTIR and XPS methods are given in the next section.

Different analytical tool gives different information about the adsorbed molecules. For instance, FTIR can give information about the orientation of adsorbed molecules as well as their packing densities [29-30, 33], while XPS can be used for determining the elemental composition, as well as the coverage density of the adsorbate on the gold surface [5-7]. Similarly, different forms of SPM have been used to characterize SAMs in various ways, ranging from monitoring the electrochemical transformation of a ferrocene-terminated alkanethiolate monolayer on gold [31-32] to chain length dependence of frictional force [33]. The technique used for the characterization of SAM at any given instance is dictated by the property to be measured.

Immobilization and Characterization of ssDNA on gold

The immobilization of single-stranded DNA (ssDNA) on gold follows the same approach as that of alkanethiols [5]. The ssDNA is always thiolated such that when a gold surface is exposed to its buffered aqueous solution, the ssDNA molecules are chemisorbed on the gold via the thiol group as shown in Figure 1.2 [5, 7-8]. Thiolated ssDNA molecules are chemisorbed on gold with Langmuir kinetics, similar to that observed for alkanethiols [6]. The molecules progressively get adsorbed on the surface until a close-packed, ordered monolayer is formed, when the packing density of the molecules reaches its limiting value.

It has been established that the presence and concentration of buffer plays a significant role in the immobilization efficiency of ssDNA [6-7]. Thus, in the absence of buffer, the ssDNA dissolved in water does not adsorb on gold, whereas the amount of adsorbed ssDNA increased exponentially as the buffer concentration increases [6-7]. Therefore, the ionic strength is important in determining the surface coverage of the ssDNA, given the fact that these molecules are negatively charged with ionizable phosphate groups. It has thus been postulated that intermolecular electrostatic repulsion between neighboring strands of DNA is minimized under high ionic strength conditions where the charged strands are better electrostatically shielded, allowing higher surface coverage of ssDNA [7].

The identity of the salts used in buffer solutions also affects the immobilization efficiency of the ssDNA [7]. Generally speaking, divalent cation buffer salts give better immobilization and hence higher surface coverage than their monovalent counterparts [7]. Consequently, it is speculated that for divalent cations to give such a high packing density for ssDNA, divalent cations may cause intermolecular or intramolecular electrostatic cross-linking of the negatively charged ssDNA. Thus, these factors can be exploited in preparing solutions to form SAMs of ssDNA on gold. For example, in order to obtain higher surface coverage for a certain ssDNA, a more concentrated buffer may be used. Also, buffers containing divalent cations may be used instead of those containing monovalent cations.

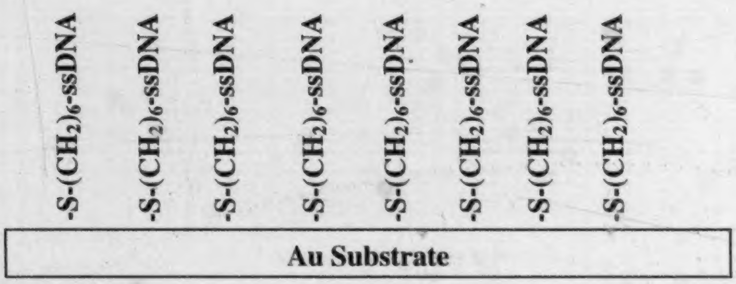


Figure 1.2: Schematic of an ssDNA immobilized on a gold substrate via the thiol group. Mixed monolayers of ssDNA and MCH can be formed by immersing the substrate in a solution of MCH, following the adsorption of ssDNA.

A large number of techniques are being used for the characterization of oligonucleotides on gold, like alkanethiols. These include FTIR [5-6, 14, 29-30, 33-38], XPS [5-7, 14, 28, 30, 34-35, 39-43], ellipsometry [14, 25, 28-29, 40, 44], atomic force microscopy (AFM) [14, 31, 34, 40], surface plasmon resonance (SPR) [5-6, 9,34, 45-49], fluorescence [5-6, 50], electrochemical methods [5, 14, 29, 39, 51-53], radiolabelling [5-7, 40, 51], contact angle measurements [14, 25, 28-29, 40, 44] and time-of-flight secondary ion mass spectrometry (ToF-SIMS) [41-42]. Of these methods, FTIR and XPS have been used complementarily for both quantitative and qualitative characterization of ssDNA [5-6]. Thus, in the work presented in this thesis, both FTIR and XPS were complementarily used for the characterization of the immobilized ssDNA. FTIR was used qualitatively, while XPS was used for both qualitative and quantitative characterizations.

The basic principle of FTIR involves the interaction of infra-red (IR) radiation with chemical compounds. This interaction causes bonds between atoms to vibrate. This vibration can be of various forms, i.e. stretching, bending, twisting, rocking and wagging, and each of these vibrational modes occurs at certain frequencies (wavenumbers) that are characteristic of the specific functional groups present in the chemical substance [54]. Molecules therefore absorb based on the type and number of functional groups present. The essential requirement for a molecule to absorb in the IR region is that the vibrations of the bonds must induce a net change in dipole moment. It is the alternating electric field produced by the changing dipole moment accompanying a vibration, which couples the molecular vibration with the oscillating electric field of the radiation. Thus, it is not necessary for the molecule to possess a permanent dipole moment for it to be IR-active [54-56]. The most useful region of IR radiation for chemical analysis is the mid-infrared region (200 - 4000cm^{-1}), and the spectrum is generated by measuring the absorption of the IR radiation by the molecule as a function of wavenumbers [54]. Wavenumbers are the units used to describe radiant energy in FTIR spectroscopy. They are proportional to the energy of vibration as well as the frequency, and modern instruments are linear in these wavenumber units [54, 56]. One advantage of FTIR is that all of the IR frequencies reach the detector simultaneously (multiplex advantage), meaning that less time is required to carry out experiments using an FTIR instrument. Most commercial instruments are equipped with a Michelson interferometer, which modulates this high frequency signal to a lower frequency that can be decoded by conventional IR detectors. Thus in FTIR, the time domain signal produced by the interferometer (interferogram), is mathematically converted to a frequency domain signal by means of Fourier transform to generate the spectrum. Once the spectrum is generated, peak assignments can be done to identify the functional groups present, with a view to identifying the compound [54].

In XPS, the sample is irradiated with soft X-ray electrons (typically 1-2keV), which cause excitation of the inner shell electrons of target atoms [57]. This induces direct emission of photoelectrons [57-58]. The kinetic energy of the emitted core electron is a function of the energy of the exciting radiation, the binding energy of the emitted electron and the work function of the instrument [58]. This kinetic energy is characteristic of the target material and measurement of the energy spectrum (counts vs binding energy) provides valuable chemical information about the top 2-20 atomic layers, depending on the material being studied [57]. Thus, the XPS spectrum gives information such as elemental composition of surfaces, chemical state identification of surfaces, composition analysis of samples, depth distribution of the composition and chemical states, as well as assessment of thin film thickness [57-58]. Peak position and peak area are used to evaluate composition, while peak shape may give information about the chemical bonds [57].

The features of these spectroscopic techniques are harnessed in the analyses carried out in chapter 2 of this thesis.

Computations

In our computational work, which is presented in chapters 3 and 4, both semiempirical PM3 (Parametric Method 3) and density functional theory (DFT) approaches were used. Semiempirical methods are based on approximations to the *ab initio* Hartree-Fock (HF) method in order to ensure that computations on large molecules are completed in reasonable CPU time. So, these methods neglect most of the electron integrals, which are evaluated in the *ab initio* HF methods [59]. PM3 method follows the Dewar's approach and so may be called Dewar-type theory. These theories treat only the valence electrons of molecules and use minimal basis set of valence Slater-type s and p orbitals. Normally, for calculations of many molecular properties, molecular orbitals are expressed in terms of basis functions. The choice of these functions always marks the start of the calculations. Thus, the success of any calculation is based on the adequate choice of basis functions. A commonly used set of basis functions is the set of Slater-type orbitals (STOs). In PM3 method, only the valence electrons of molecular systems are considered and minimal basis set is used. A minimal basis set consists of one STO for each inner-shell and valence-shell atomic orbital (AO) of each atom. The Dewar-type theories are parameterized to give good values of the 25°C gas phase standard enthalpy of formation [59]. PM3 is similar to another Dewar-type theory called AM1 (Austin Model 1). In fact, PM3 is an offshoot of AM1. PM3 however differs from AM1 because it is a reparameterization of the AM1 theory and so uses different values of parameters. Also, PM3 method derived its parameters by comparing much larger number and wider variety of experimental versus molecular properties [60], and so may be relatively more accurate than AM1 method. It must be pointed out that, since PM3 method is not parameterized for total energy, the total energy calculations carried out with this method in chapter 4 are only approximate, and cannot be as accurate as the DFT calculations.

The DFT method has the advantage of including electron correlation in a calculation, an effect that is a major shortcoming of the HF method, and it takes about the same time as a HF calculation [59]. DFT is based on the assumption that the ground state molecular energy and other molecular properties can be determined from the ground state electron density. These properties are evaluated by means of exchange-correlation functionals, which are used to express the electron density. Within DFT, other approximations are made, which are employed for molecular properties calculations. The local-density approximation (LDA) assumes a slowly varying electron density as a function of position. A similar but slightly better approximation is the local-spin-density approximation (LSDA), where electrons of opposite spins occupy different spatial orbitals, as opposed to the LDA where such electrons are paired in the same spatial orbital. These two approximations may be inappropriate for many molecular systems because the electron density is not a slowly varying function of position. The functionals that are based on these approximations are called local functionals. The gradient-corrected approximation (also called generalized gradient approximation GGA) was formulated to account for the electron inhomogeneity inherent in molecular systems by including the gradient of the density in its formulation. In addition to the local and gradient-corrected functionals which are based on the above two approximations, there are hybrid functionals which mix together the HF, local and gradient terms in their exchange-correlation parts. Both gradient-corrected and hybrid functionals are known to perform better than local functionals for many molecular systems [59]. The DFT calculations presented in chapters 3 and 4 use all the three types of functionals.

Two basis sets were used in our DFT calculations. These are the medium basis set 6-31G* and the large basis set 6-31G**, both of which are called split-valence basis set. Split-valence basis sets use two or more STOs for each valence AO and only one STO for each inner-shell AO. Each of these may or may not include polarization functions, these are used to account for distortion of atomic orbitals, which occur when orbitals overlap. However, Gaussian-type functions (GTFs) are used instead of STOs for the atomic orbitals to speed up molecular integral evaluation [59]. Thus, the STOs are usually expressed in terms of the GTFs. In addition, instead of using individual GTFs as basis functions, they may be taken as normalized linear combination of a few GTFs. The old GTFs are called primitive Gaussians; while the new normalized ones are called contracted Gaussian-type functions (CGTF). The idea behind this practice is that, by using contracted Gaussians instead of primitive Gaussians, a great deal of savings can be achieved in computation time with little loss in accuracy [59]. The 6-31G* basis set uses six primitives in each inner-shell CGTF and represents each valence-shell AO by one CGTF with three primitives and another CGTF with one primitive, hence the 6-31 notation. Also the presence of one asterisk indicates that one d-type polarization function (with six or five components) are added to each second and third row atom, and f-type polarization functions on each atom from Sc-Zn. The 6-31G** basis set is similar to the 6-31G* basis set, but in addition to all the specifications stated above for 6-31G*, adds a set of one p-type polarization function (with three components) on each H and He atoms.

The DFT calculations were carried out using tight convergence criteria. This is a convergence-related option in Gaussian 03 [61], used to tighten the cutoffs on forces and step size that are used to determine convergence. This is especially necessary for molecular systems with very small force constants to ensure adequate convergence and reliability of computed vibrational frequencies. Ultrafine numerical integration pruned grids were also used for all our DFT calculations, as recommended in the software manual. Pruned grids are grids that have been optimized to use the minimal number of points required to achieve a given level of accuracy [61].

Research Objectives

Fabrication and characterization of molecular architectures of different forms are now developed to a stage where it can be harnessed to gain insights into the properties of a variety of systems, ranging from alkanethiols, polymers and biomolecules. Different approaches can be used to form an array of the compound of interest on a substrate, which can then be appropriately characterized using different analytical tools.

The formation of SAMs of ssDNA on gold surface forms the basis for a range of important biotechnology applications. These include DNA microarrays, biosensors, gene mapping and nanoscience applications. However, little is known about the immobilization mechanism, surface density and the film structure of these molecules on the surface. The goal of my research in this respect is to explore the conventional immobilization technique to form an assembly of a different and unique ssDNA on gold surface and use the two complementary surface analytical tools (FTIR and XPS) to characterize the adsorbed molecules. It is hoped that a standard immobilization procedure for ssDNA will be established, in terms of concentration of adsorbates, buffer type and concentration, and pH range. Also, the results of these characterizations will give us a better understanding of the immobilization mechanism, surface coverage as well as the film structure of this system. Thus, chapter 2 of this thesis is dedicated to the immobilization and characterization of different oligonucleotides on gold.

In chapter 4, I used computational approach to explore the properties of a similar system (polyynes), which can be equally used as building blocks for molecular architectures like ssDNA. chapter 3 deals with the selection of a suitable density functional theory (DFT) functionals that will be appropriate for the bending and stretching calculations to be carried out in chapter 4. Fifteen different functionals including local, gradient-corrected and hybrid functionals were tested on a model molecule with a view to selecting the functionals of best overall performance. This selected functional was then used for subsequent calculations.

In chapter 4, the unique properties of polyynes were examined using PM3 and DFT methods. These molecules are believed to have different potential applications in high technology industries. A good understanding of the properties of these compounds on the molecular scale will enhance our capabilities in designing devices with desirable properties and applications. Thus, we used a computational approach to probe the possible mechanical and electronic properties of these molecules. We used various levels of theory, from semi-empirical to *ab initio* / DFT for these calculations. We monitored how the total energy changes with respect to bending, as well as the variation of total energy with stretching coordinates. The bending calculations are especially relevant as some measurements can be made with an atomic force microscope (AFM). Thus, the calculations allowed us to make reasonable predictions of how these molecules might respond to deformations that can be caused by AFM tip. The responses of these molecules in terms of how the relative force constants with which they might push on an AFM tip or be bent by the tip were found.

References

1. C. A. Reed, *Acc. Chem. Res.* **38,** 215 (2005).
2. P. D. W. Boyd and C. A. *Reed, Acc. Chem.* Res. **38,** 235 (2005).
3. S. I. Nishikiori, H. Yoshikawa, Y. Sano and T. Iwamoto, *Acc. Chem. Res.* **38,** 227 (2005).
4. D. M. Sarno, et al. *Langmuir,* **16,** 6191 (2000).
5. D. Y. Petrovykh, H. Kimura-Suda, M. J. Tarlov and L. J. Whitman, *Langmuir,* **20,** 429 (2004).
6. D. Y. Petrovykh, H. Kimura-Suda, L. J. Whitman and M. J. Tarlov, *J. Am. Chem. Soc.* **125,** 5219 (2003).
7. T. M. Herne and M. J. Tarlov, *J. Am. Chem. Soc.* **119,** 8916 (1997).
8. M. J. Tarlov and A. B. Steel, *Biomolecular Films: Design, Function, and Applications;* J. F. Rusling, Ed.; Mercel Dekker: New York, 2003; **Vol. 111.**
9. A. W. Peterson, L. K. Wolf and R. M. Georgiadis, *J. Am. Chem. Soc.* **124,** 14601 (2002).
10. U. Molder, P. Burk, I. A. Koppel, *Int. J. Quant. Chem.* **82,** 73 (2001).
11. F. Coat, C. Lapinte, *Organometallics,* **15,** 477 (1996).
12. H. S. Nalwa, J. Mukai, A. Kakuta, *J. Phys. Chem.* **99,** 10766 (1995).
13. D. W. Rogers, et al. *J. Org. Chem.* **69,** 7143 (2004).

14. A. Ulman, *Chem. Rev.* **96,** 1533 (1996).
15. A. D. Slepkov, et al. *J. Chem. Phys.* **120,** 6807 (2004).
16. M. Tsuji, S. Kuboyama, T. Matsuzaki, T. Tsuji, *Carbon,* **41,** 2141 (2003).
17. T. D. Poulsen, et al. *J. Chem. Phys.* **114,** 5917 (2001).
18. S. Eisler, et al. *J. Am. Chem. Soc.* **127,** 2666 (2005).
19. M. A. Heuft, S. K. Collins, G. P. A. Yap, A. G. Fallis, *Org. Lett.* **3,** 2883 (2001).
20. J. L. Toto,T. T. Toto, C. P. de Melo, *Chem. Phys. Lett.* **245,** 660 (1995).
21. C. K. Chiang, et al. *Phys. Rev. Lett.* **39,** 1098 (1977).
22. A. D. Slepkov, et al. *J. Chem. Phys.* **116,** 3834 (2002) .
23. F. Coat, C. Lapinte, *Organometallics,* **15,** 477 (1996).
24. Q. Fan, G. V. Pfeiffer, *Chem. Phys. Lett.* **162,** 472 (1989).
25. D. K. Schwartz, *Annu. Rev. Phys. Chem.* **52,** 107 (2001).
26. R. G. Nuzzo and D. L. Allara, *J. Am. Chem. Soc.* **105,** 4481 (1983).
27. R. G. Nuzzo, F. A. Fusco and D. L. Allara, *J. Am. Chem. Soc.* **109,** 2358 (1987).
28. C. D. Bain, et al. *J. Am. Chem. Soc.* **111,** 321 (1989).
29. O. Chailapakul, L. Sun, C. Xu and R. M. Crooks, *J. Am. Chem. Soc.* **115,** 12459 (1993).
30. A. Ulman, *An Introduction to Ultrathin Organic Films,* Academic Press: Boston, 1991.
31. J. -B. D. Green, M. T. McDermott and M. D. Porter, *J. Phys. Chem.* **100,** 13342 (1996).
32. L. A. Bottomley, *Anal. Chem.* **70,** 425R (1998).
33. M. T. McDermott, J. -B. D. Green and M. D. Porter, *Langmuir,* **13,** 2504 (1997).
34. E. A. Smith, et al. *Langmuir,* **17,** 2502 (2001).
35. A. V. Saprigin, et al. *Surf. Int. Anal.* **36,** 24 (2004).
36. L. M. Demers, et al. *J. Am. Chem. Soc.* **124,** 11248 (2002).
37. H. Kimura-Suda, D. Y. Petrovykh, M. J. Tarlov and L. J. Whitman, *J. Am. Chem. Soc.* **125,** 9014 (2003).

38. D. Y. Petrovykh, et al. *J. Am. Chem. Soc.* **128,** 2 (2006).

39. J. Wang, G. Rivas, M. Jiang and X. Zhang, *Langmuir,* 15, 6541 (1999).

40. S. G. Ray, H. Cohen, R. Naaman and Y. Rabin, *J. Am. Chem. Soc.* **127,** 17138 (2005).

41. C. J. May, H. E. Canavan and D. G. Castner, *Anal. Chem.* **76,** 1114 (2004).

42. C. Y. Lee, H. E. Canavan, L. J. Gamble and D. G. Castner, *Langmuir,* **21,** 5134 (2005).

43. Y. D. Zhao, et al. *Anal. Chim. Acta,* **388,** 93, (1999).

44. T. Aqua, R. Naaman and S. S. Daube, *Langmuir,* **19,** 10573 (2003).

45. A. W. Peterson, R. J. Heaton and R. M. Georgiadis, *Nucl. Acid. Res.* **29,** 5163 (2001).

46. K. A. Peterlinz, R. M. Goergiadis, T. M. Herne and M. J. Tarlov, J *J. Am. Chem. Soc.* **119,** 3401 (1997).

47. H. J. Lee, T. T. Goodrich and R. M. Corn, *Anal. Chem.* **73,** 5525 (2001).

48. M. B. Esch, L. E. Locascio, M. J. Tarlov and R. A. Durst, *Anal. Chem.* **73,** 2952 (2001).

49. L. K. Wolf, Y. Gao, and R. M. Georgiadis, *Langmuir,* **20,** 3357 (2004).

50. U. Rant, et al. *Langmuir,* **20,** 10086 (2004).

51. A. B. Steel, R. L. Levicky, T. M. Herne and M. J. Tarlov, *Biophys. J.* 79, 975 (2000).

52. A. B. Steel, T. M. Herne and M. J. Tarlov, *Anal. Chem.* **70,** 4670 (1998).

53. A. B. Steel, T. M. Herne and M. J. Tarlov, *Bioconj. Chem.* **10,** 419 (1999).

54. D. A. Skoog, F. J. Holer and T. A. Nieman, Principles of Instrumental Analysis, Fifth Edition; Thomson Learning, 1998.

55. I. L. Finar, Organic Chemistry Volume 1, The Fundamental Principles, Sixth Edition; Longman Group Limited.

56. R. M. Silverstein, G. C. Bassler and T. C. Morill, *Spectrometric Identification of Organic Compounds,* Fifth edition; John Wiley & Sons, Inc. New York, 1991.

57. http://www.ualberta.ca/ACSES/ACSES3/Techniques/XPS.htm

58. J. B. Lumsden, ASM Handbook, v.10, Materials Characterization, ASM International, (1990). Also available online at http://www.library.ualberta.ca/databases/databaseinfo/index.cfm?ID=3272
59. I. N. Levine, *Quantum Chemistry*, Prentice Hall (2000).
60. *HyperChem Computational Chemistry, Practical Guide, Theory and Methods,* Hypercube, Inc., 1996.
61. http://www.gaussian.com/g_ur/k_opt.htm

CHAPTER 2

Spectroscopic Characterization of Oligonucleotides Immobilized on Gold

Introduction

Single-stranded DNA (ssDNA) immobilized on gold surface forms the basis for a number of important biotechnology applications, such as DNA microarrays, biosensors, gene mapping and nanoscience applications [1-5]. Despite these excellent possible applications, little quantitative information can be obtained about the immobilization mechanism and the film structure of the molecules on the surface. For instance, accurate measurement of surface density, which is important for determination of immobilization and hybridization efficiency, is difficult to come by [1-2, 4].

An array of different methods has been used over the years for different forms of characterization of oligonucleotides on surfaces, some of which are used for quantitative characterization. This includes electrochemical methods [1, 6-9], fluorescence [1-2, 10], neutron reflectivity [11], surface plasmon resonance (SPR) [1-2, 5, 12-17], atomic force microscopy (AFM) [17-18], contact angle measurement technique [18-19], ellipsometry [18-19], radiolabelling [1-3, 6, 18], Time-of- flight Secondary Ion Mass Spectrometry (ToF-SIMS) [20-21], Fourier Transform Infrared Spectroscopy (FTIR) [1-2, 17, 22-25] and X-ray Photoelectron Spectroscopy (XPS) [1-3, 9, 17-18, 20-22, 26].

Of these methods, radiolabelling has been used extensively, relatively speaking, for quantification. However, the use of radiolabels is discouraged on grounds of health hazards. Also, label-free methods like ellipsometry and SPR can provide real-time measurements, but they are incapable of screening out non-specific adsorption, which complicates the results. XPS is a label-free method that can provide quantitative and qualitative information about DNA films adsorbed on the surface, it is however incapable of screening out non-specific adsorption of these molecules. If an adequate model is used for the films and realistic assumptions are made, it can give accurate quantitative information about the films [1]. It can also be employed to give chemically specific (qualitative) information about the DNA films.

FTIR is another method that can be used to qualitatively characterize DNA films adsorbed on gold using the characteristic peak positions of specific functional groups. With this technique, non-specific adsorption of ssDNA can be identified based on the characteristic peak positions of chemisorbed molecules. When the peak areas are appropriately found, it can also be harnessed to give quantitative information about the adsorbed DNA films on the surface. Thus, these two methods can complement other analytical techniques that are being used for characterization of DNA adsorbed onto gold [1, 3].

In this chapter, we describe the immobilization of ssDNA (thiolated TT and non-thiolated AA) on gold surfaces. We also present the results of characterization experiments performed on the ssDNA bound to the gold surface. We used XPS and FTIR as complementary techniques to characterize these immobilized molecules. Immobilization of aqueous solutions of thiolated ssDNA on the gold surface occurs by chemisorption on the gold surface via the thiol group [1, 3-4]. This is similar to the formation of self-assembled monolayers (SAMs) of alkanethiols on gold [1]. Non-thiolated ssDNA samples do not have this property of being adsorbed on the gold surface via thiol group, the molecules can only be adsorbed non-specifically, probably via the nucleotide bases [3]. After the immobilization, the surface density of the adsorbed molecules can be controlled by backfilling with 6-mercapto-1-hexanol (MCH), which minimizes non-specific adsorption by displacing non-specifically adsorbed molecules off the surface [3]. This spacer molecule has been extensively used for this purpose in the immobilization of aqueous solutions of DNA on gold. We characterized homo-oligonucleotides (i.e. pure TT and AA nucleotides of 20 bases long), which are relatively simple and have specific spectral signatures, which make the interpretation of the resulting data less cumbersome. We obtained qualitative information from the XPS results, while our FTIR results revealed the unique spectral features of the nucleotides. The results give information on the film structure and interaction of the molecules with the surface. We also show how the XPS and FTIR results complement each other.

Experimental

Materials and Reagents

All solutions were prepared with deionized water (~18MΩ.cm) from Barnstead Nanopure (Dubuque, IA) purification system. Thiolated and non-thiolated single-stranded oligonucleotides (ssDNA) >85% were purchased from IDT (Coraville, IA) with HPLC purification and were used as received without further modifications. NaCl (99.0%)and NaOH (97.0%) were obtained from EM Science (Gibbstown, NJ), 6-mercapto-1-hexanol (MCH) 97%, was obtained from Aldrich (Milwaukee, WI) Tris [Tris (Hydroxymethyl) Aminomethane] was obtained from ICN Biomedicals, Inc. (Aurora, OI), while EDTA (99.4%) was from BDH Inc. (Toronto, ON). HCl (36.5%) was from Fisher (Nepean, ON), while H_2O_2 (30%) and H_2SO_4 (95.0%) were both purchased from EMD Chemicals, Inc. (Gibbstown, NJ). Thermal evaporator used for the deposition of Cr/Au substrates was from Torr International Inc. (New Winsdor, NY), while UV ozone cleaner (Jelight Company Inc. Irvine, CA) was used for cleaning Au-coated glass slides for FTIR experiments. 10TE buffer (100mM Tris and 10mM EDTA solution with pH adjusted to 7.50 with 2M HCl) were used to prepare buffer solutions. 1TE-NaCl buffer (a solution containing 1M each of TE and NaCl) was used for all experiments. Bulk ssDNA solution was prepared by dissolving the original "as received" samples in 1TE-NaCl buffer and confirming its concentration by UV absorption measurements. For example, the absorbance reading for sample TT (diluted ten times) of molar extinction coefficient 162,600 L/mole.cm was 0.565 a.u. in a sample cell of 1cm. This gave a molar concentration of 34.74 x 10^{-6}M after multiplying by a dilution factor of 10. A 720μL aliquot of this bulk solution was then diluted to 25ml with 1TE-NaCl buffer to give a 1μM solution of this sample. The resulting 1μM DNA solution was used for immobilization experiments. Piranha solution, consisting ~30% H_2O_2 and ~70% H_2SO_4 (by volume) was used for cleaning silicon wafers and gold substrates.

Preparation of ssDNA Films

Gold films on single-crystal Si(100) wafers were used as substrates. Typically, prior to deposition, the wafers were cleaned in piranha solutions for about 1 hour. These were then rinsed thoroughly with deionized water and then blown dry in a stream of nitrogen gas. A Cr adhesion layer (15.7nm) was deposited by thermal evaporation, followed by a 200nm of Au. Each substrate was again cleaned in piranha solution for about 75 minutes, rinsed with deionized water and dried with nitrogen.

The ssDNA SAMs were typically prepared by immersing clean gold substrates (~1cm^2) in 1μM DNA solution (prepared from the bulk sample solution as explained above) at ambient temperature and left overnight (about 18 hours). Those backfilled with MCH were then immersed in 1.0mM of this solution, while the remaining ones were rinsed and dried. The backfilled ones were later removed from MCH, rinsed and dried.

Samples for FTIR measurements were prepared by immersing glass slides coated with 150nm Au/5nm Cr (which were cleaned in piranha solution before evaporation as was done for the silicon wafers, and just before use, rinsed with EtOH and cleaned in UV ozone cleaner for 10 minutes) in 1μM ssDNA solutions overnight. They were then removed, rinsed and dried before measurements. Tables 2.1 and 2.2 give the description of the samples used in this work.

Name	**Sequence**
TT	**5'HS-S-C6-TTT TTT TTT TTT TTT TTT TT-3'**
AA	**5'-AAA AAA AAA AAA AAA AAA AA-3'**

Table 2.1: The base sequence of the two main ssDNA samples used in this work. Sample TT is thiolated while AA is non-thiolated, each of which is 20 nucleotides long. A 1μM aqueous solution in 1TE-NaCl buffer of each was used in all experiments.

Sample	**Description**
TT	Thiolated DNA containing T bases.
TT/MCH	Sample TT backfilled with 6-mercapto-1-hexanol (MCH).
AA	Non-thiolated DNA containing A bases
AA/MCH	Sample AA backfilled with 6-mercapto-1-hexanol (MCH).
Control	Control sample containing only the buffer solution without DNA.
MCH	6-mercapto-1-hexanol [$HO(CH_2)_6SH$]

Table 2.2: A brief description of the samples used. Backfilled samples were treated with MCH for about 1hour after spending about 18hours in DNA solutions.

XPS Measurements

XPS measurements were performed using a commercial high-resolution XPS system (Axis 165 X-ray Photoelectron Spectrometer from Kratos Analytical, Chestnut Ridge, NY) equipped with dual (Mg and Al) and monochromatic Al X-ray sources, a 165mm mean radius hemispherical analyzer and an eight channeltron detection system for excellent energy resolution and sensitivity at small analysis areas [27]. The spot size used for the analyses is 700x400um and data was collected from only one spot on each sample. The absolute peak intensities indicated are based on the count recorded by the analyzer as calculated by the acquisition software. Two types of scans were acquired: survey scans from 0 to 1000 eV BE and 160 pass energy and high-resolution scans with 12 eV windows and 20 eV pass energy. Pass energy is the kinetic energy an electron must possess in order to pass through the analyzer and be detected. The high-resolution scans were acquired for Au4f, O 1s, C 1s, N 1s, P 2p and S 2p regions. These scans were used to determine stochiometry and coverage for the DNA films. Peak areas were obtained from the high-resolution spectra by transferring the raw data in ASCII format to excel spreadsheet for processing and then to Igor Pro, the software used for calculating the peak areas.

FTIR Measurements

FTIR absorption spectra were acquired with an ATI Mattson Infinity Series FTIR Spectrometer (Madison, WI) equipped with a Mercury-Cadmium-Telluride (MCT) detector cooled with liquid nitrogen. Spectra (800 - 4000cm^{-1}) were collected with 500 scans at 2 cm^{-1} resolution. A deuterated octadecanethiol sample was used as the background.

Results and Discussion

XPS Results

XPS was used to obtain high-resolution spectra for all the principal elements in immobilized ssDNA film (N, C, O and P). High-resolution spectra were also obtained for the thiol group to gain information about the immobilization of the ssDNA on gold [1]. These high-resolution spectra were obtained for all the five samples including the control.

Before treating the results in detail, it is pertinent to briefly discuss the general properties of a ssDNA film on gold. The presence of N [1-3, 28], and to a better extent, N and P together, is an excellent indicator of adsorbed DNA [1-2, 26] because their presence is typically unaffected by surface contamination during sample preparation and handling [1, 2]. C and O on the other hand are unsuitable as markers of adsorbed DNA because they are prone to contamination [26]. In addition, both thymine and adenine deoxyribonucleotides contain more N atoms than P, and N has a higher XPS cross-section [2], N thus provides a more reliable reference for composition measurements [2] and determination of DNA coverage [1]. Therefore, N will be used in this report as the frame of reference for all samples.

Signal intensities and Peak areas

Table 2.3 shows the peak areas of the different elements of the samples considered and that of gold substrate. The P2p and S2p peaks are for the most part weak and can hardly be used for quantitation. The O1s and C1s peaks are not to be reckoned with, since these elements are prone to contamination.

Sample	O1s	N1s	C1s	S2p	Au4f
TT	3.56	0.28	0.91	0.02	44.32
TT/MCH	3.45	0.17	0.85	0.04	44.22
AA	0.99	0.77	0.85	0.04	107.34
AA/MCH	0.82	0.74	0.83	0.03	111.15
Control	3.90	0.17	0.82	0.02	41.75

Table 2.3: Peak areas (in millions) of the major XPS signals of the five samples including the control. Peak areas are in units of counts-electronvolt (c.eV)

One important observation made from the table is that contrary to expectation, the N 1s peak areas for the non-thiolated ssDNA samples AA and AA/MCH are both significantly higher than both TT and TT/MCH, which are thiolated ssDNA samples.

This observation may be seen as one that raises some ambiguity on the validity of the data, however, the observation can be rationalized based on some similar results reported in the literature. Non-thiolated ssDNA molecules can adsorb onto a gold surface by non-specific adsorption through the nucleotide bases or some other functionality of the ssDNA [3].

Significant amounts of non-thiolated ssDNA (especially oligo poly-A) have been reported to adsorb onto gold surface, suggesting that even in the absence of thiol group, ssDNA will adsorb on the surface [3, 16, 23-24, 28]. For instance, Kimura-Suda in his base-dependent competitive adsorption studies of ssDNA on gold [24], found that even when lower concentration of non-thiolated adenine nucleotides (dA_5) are present together with higher concentration of thiolated thymine nucleotides (dT_{25}), dA is still preferentially adsorbed on the gold surface. Wolf also found a similar trend in the relative adsorption of poly (dT) and poly (dA) [16]. In the same vein, Rapino in his computational modeling studies on the stability and motion of DNA nucleobases on gold, found single and pairs of adenine bases to have stronger affinity for gold than their thymine counterparts [28], an observation also reported by Demers [23] in his thermal desorption and binding studies. The results of another similar study led to the speculation that, even for thiolated ssDNA, the nucleotide side chains play a role in the adsorption, such that initially, the nucleotide side chains may be interacting directly with the surface, before the molecule reorganizes itself on the surface when the thiol group would become the primary anchor [3].

All these observations may lead us to speculate here that sample AA interacts more strongly with the gold surface than the thiolated TT, hence, the larger peak area, in agreement with similar reports just cited. Thus, when similar concentrations of thiolated TT and non-thiolated AA are used for immobilization, AA may have a larger peak area due its stronger affinity for gold. Or, it may be that certain contaminants coadsorbed with sample TT on the gold surface, thereby burying some of the N atoms from the DNA molecules, resulting in the observed weak N1s signal. In any case, this conclusion has to be validated by further experiments involving many samples similarly prepared.

The N 1s peak for sample TT is larger than that of TT/MCH, which could mean that expectedly, the MCH indeed displaced some immobilized ssDNA molecules. The N 1s peak area of the TT/MCH is about 62% that of sample TT. However, the N 1s peak area for the AA/MCH sample is about 96% that of sample AA. An N 1s signal was also observed for the control sample, an observation that has been reported by some authors [29]. This is however contrary to a report that bare gold samples exposed to buffer solutions containing no DNA exhibit no XPS-detectable nitrogen [3], probably because a different buffer was used (phosphate buffer vs 1TE-NaCl buffer used here). It is also noted that the signal intensities of all elements in the control sample are similar to those in some or all of the other samples. This also explains why it is necessary that the data presented here be validated by carrying out experiments on many similarly prepared samples.

Considering the Au4f signal intensities, it can be seen that the signal is more strongly attenuated for samples TT and TT/MCH relative to the non-thiolated AA samples. This indicates that the films of the thiolated samples are thicker than the non-thiolated ones, even though, the AA samples have higher N1s signal intensities. This further supports the above argument that it is likely that TT coadsorbed with contaminants, which makes the Au signal to be so attenuated and the N1s signal so weak. In addition, it was observed that the C1s and S2p signals are essentially the same for all samples.

Principal elements peak positions

The high-resolution XPS spectra of the principal elements in an immobilized ssDNA film are shown in figures 2.1 to 2.5. The O 1s peaks of 530.3 eV and 530.4 eV found for samples TT and TT/MCH are slightly lower than reported values for similar films (531.7-533.3 eV). Peak positions for AA and AA/MCH however, fall in the range already reported [1]. The locations of the C 1s peaks of all the samples are also in close agreement with reported literature values [1, 20-21], though the C 1s peak position of sample TT is slightly shifted to a higher energy relative to others.

The principal N 1s core-level peaks for TT and TT/MCH in figure 2.3 have BE values that are consistent with published results for thymine nucleotides [1-2]. The

TT samples have their two nitrogen atoms in similar chemical environments, thus a simple N 1s spectrum is expected [1]. The fact that samples AA and AA/MCH have slightly lower BE values for N 1s peak can be ascribed to their chemisorption on the gold surface [1-2], a phenomenon that has also been observed even for thiolated thymine nucleotides. This observation probably explains why the signal intensities of the non-thiolated AA samples are higher than those of the thiolated TT samples. This is further explained in the FTIR results section. It could also be speculated that sample TT is coadsorbed on the surface with contaminants, thereby lowering the N1s signal for this sample. This may be true, considering the signal intensities of C1s and N1s in relation to Au4f signal for samples TT, TT/MCH, AA and AA/MCH in Table 2.3.

The P 2p signal is weak for the samples except for AA and AA/MCH where the signal is a bit larger. These samples have their peaks located at 133.3 eV which is characteristic of nucleotides. The relatively large signal observed for these non-thiolated ssDNA samples relative to the thiolated ones is another indication that AA is non-specifically adsorbed on the Au surface and probably less of TT is adsorbed as indicated by the N1s signals in Table 2.3, or it is coadsorbed on the surface with contaminants which somehow buries the ssDNA molecules, thereby making the P2p signal very weak compared to AA and AA/MCH. This assumption however needs to be validated by carrying out further experiments.

The S 2p signal is extremely difficult to observe perhaps due to its very low relative concentration (assumed theoretical S/P ratio is 1/20 for our ssDNA samples since each of them is 20-bases long). Also, since it is believed that thiolated ssDNA are anchored onto the gold via thiol groups, then the sulphur signal would be strongly attenuated by the DNA film [1]. However, sample AA/MCH has an observable S 2p signal located at 162.1 eV, in excellent agreement with values attributed to thiol-gold bond in alkanethiols SAMs [1, 30-33]. But why does the non-thiolated AA backfilled with MCH (i.e. AA/MCH) have such large signal compared to other samples? This may be because the sample originally containing only non-thiolated AA was chemisorbed on the surface and when backfilled with MCH, the non-specifically adsorbed molecules are displaced (as typical of MCH action) and the S 2p signal seen in this sample is from MCH (and of course not from AA). Also, since MCH is a much shorter molecule than ssDNA, the sulphur signal is visible and not attenuated by the alcohol film. In addition, TT is believed to be mostly anchored through thiol groups, hence only relatively small number of molecules was being displaced by MCH in this case, and so the S atom is buried by the overlying ssDNA molecule, hence the S2p signal is very weak. It may perhaps be due to the fact that there are other materials which coadsorbed on the surface with this ssDNA sample, which caused its S2p signal to be very weak. Thus, the MCH cannot displace this contaminant and so the S atom remained buried under this film of contaminant, thereby making the signal very weak.

A similar report of weak S 2p signal which is difficult to observe has been made in the literature, in which case a clean signal was only observed after 30 minutes of immobilization, while longer times of immobilization gave very weak signals [1].

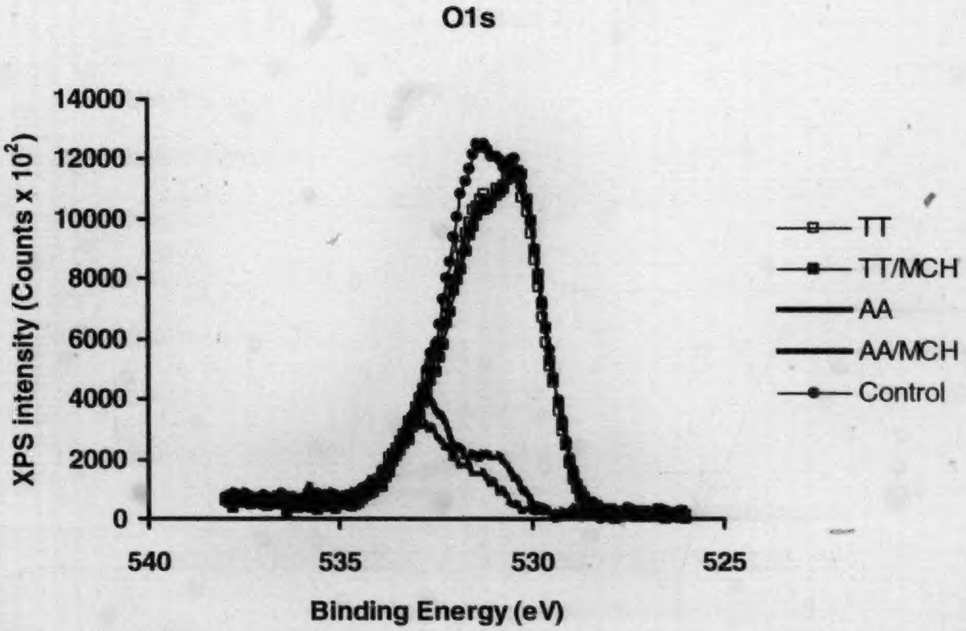


Figure 2.1: High-resolution XPS spectra of O1s peak region for the five samples. The O 1s peak is in each case located at 530.3 eV, 530.4 eV, 531.2 eV, 532.9 eV and 533.1 eV for samples TT, TT/MCH, Control, AA and AA/MCH respectively.

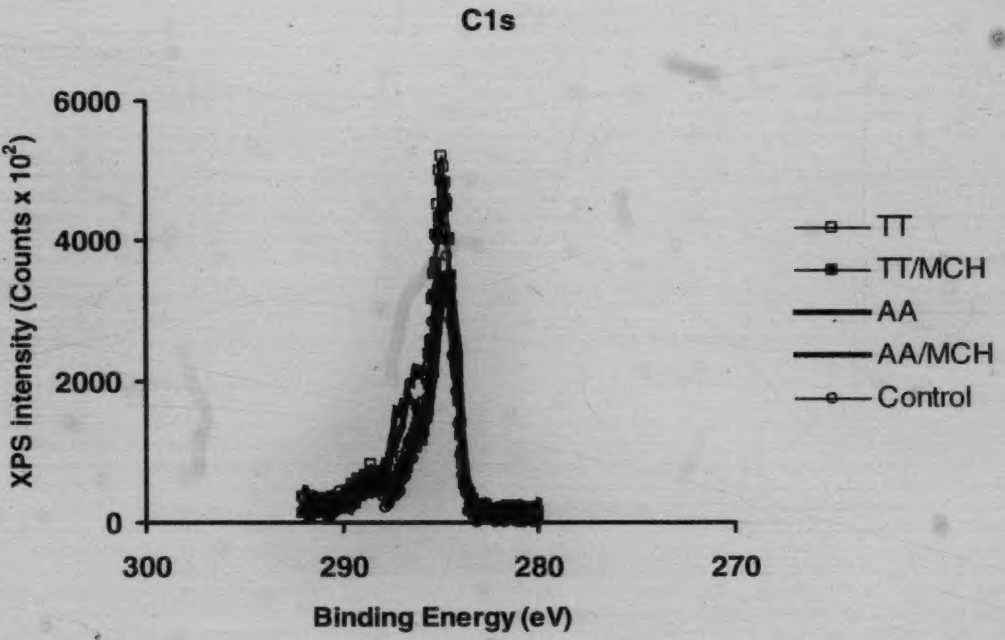
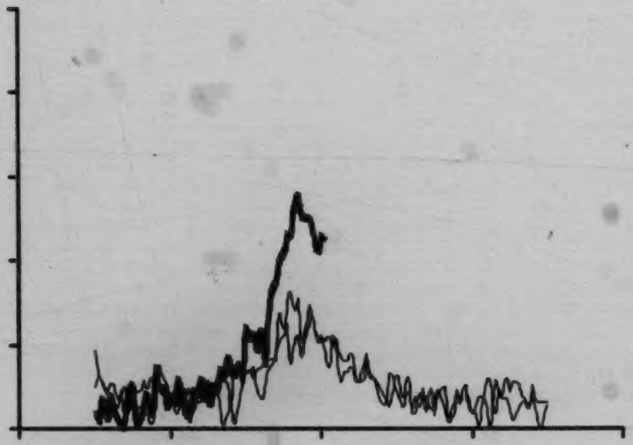


Figure 2.2: High-resolution XPS spectra of C1s peak region for the five samples. The C 1s peak of each sample is located at about 284.8 eV except for sample TT whose peak is found at 286.9 eV.



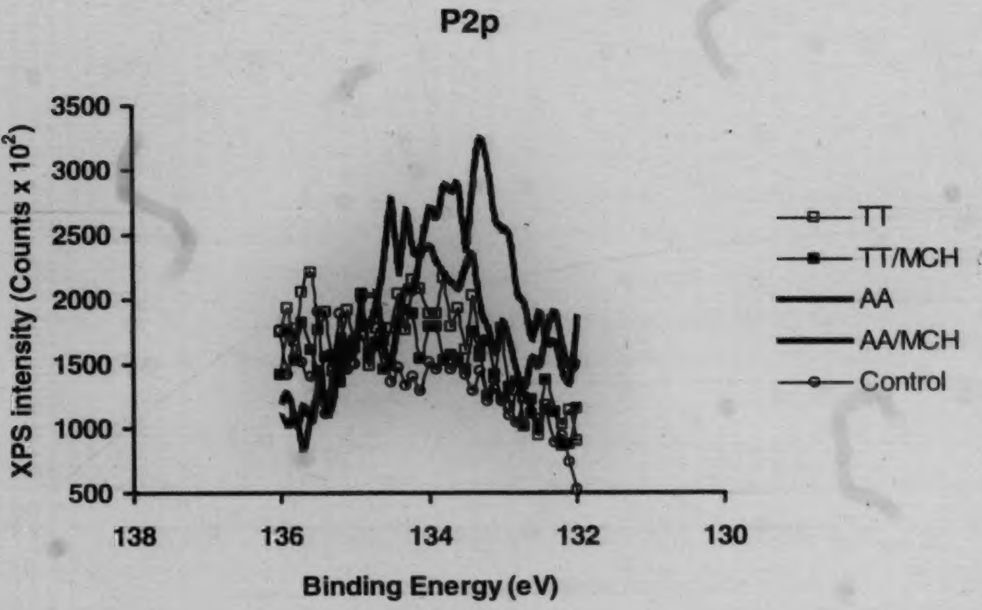


Figure 2.4: High-resolution XPS spectra of P2p peak region for the five samples. The P 2p signal is weak for all except AA and AA/MCH whose peaks are located at 133.3 eV.

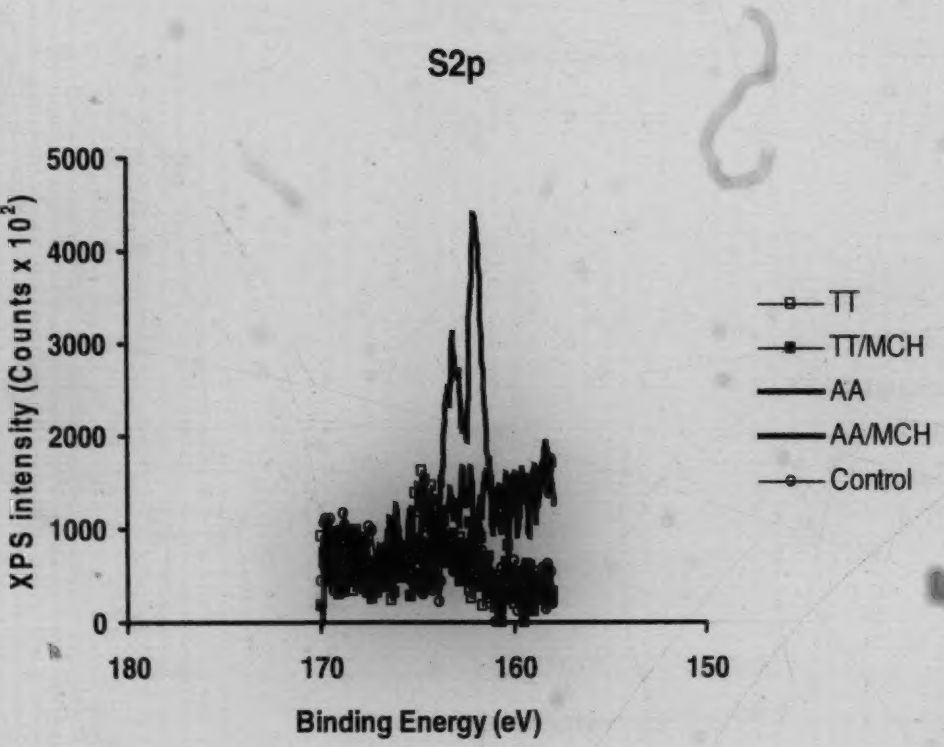


Figure 2.5: High-resolution XPS spectra of S2p peak region for the five samples. The S 2p signals are weak except for sample AA/MCH whose peak is located at 162.1 eV typical of sulphur bound to a gold surface.

Relative atomic density and DNA coverage

Though the data presented in this work have to be validated by carrying out further replicate experiments, it may be necessary to have an idea of the DNA film thickness and relative atomic density, using N as the reference atom. The procedure used here is based on the exponential attenuation and standard overlayer XPS formalism of Dmitri [1-2], which accurately accounts for the signal attenuation in the DNA film. This model uses effective attenuation length (EAL) which is calculated using the freely supplied National Institute of Standards and Technology (NIST) Standard Reference Database 82 (SRD-82) [34] for electrons in the film. The model takes into consideration the fact that signal attenuation in the DNA film is significant for photoelectrons from both the substrate and the overlayer, and that both effects must be properly accounted for in the analysis [1]. The steps are as follows:

- the film thickness *t* is determined using equation 2.1
- the obtained film thickness is then used to calculate relative atomic density, based on equation 2.2
- finally, the relative and absolute elemental coverage (DNA coverage) values are calculated based on film thickness and relative atomic density determined in the first two steps according to equations 2.3 to 2.5 below.

The following equations are used to determine the above parameters:

$$I_{Au} = I^{o}_{Au} \exp\left[-t/L_{Au}\right]$$

$$\text{Then, film thickness } t = \left[\ell n\left(I_{Au} / I^{o}_{Au}\right)\right]\left(-L_{Au}\right) \tag{2.1}$$

where I_{Au} is the observed intensity (peak area) of the gold substrate signal

I^{o}_{Au} is the intensity (peak area) of a clean gold substrate (16.35 x 10^{7} c.eV)

t is the film thickness to be determined

L_{Au} is called the "average practical EAL" (PEAL) for electrons from gold in DNA overlayer

Relative atomic density:

$$N_N / N_{Au} = I_N / I_{Au} (T_{Au} \sigma_{Au} L^{Q}{}_{Au} / T_N \sigma_N L^{Q}{}_N) \exp(-t / L_{Au}) / 1 - \exp(-t / L_N) \qquad \textbf{(2.2)}$$

where N_N / N_{Au} is the relative atomic density for N

I_N and I_{Au} are the signal intensities for N and Au

T_{Au} and T_N are the analyzer transmission function for N and Au

σ_{Au} and σ_N are the total photoelectric cross section for N and Au

$L^Q{}_{Au}$ and $L^Q{}_N$ are the effective attenuation length (EALs) for quantitative analysis (QEAL) for N and Au

L_N is the "average practical EAL" for electrons from DNA in the DNA overlayer

For DNA coverage, we have

$$\theta_N = N_N t \qquad \textbf{(2.3)}$$

where θ_N is the number of N atoms per unit area of film

N_N atomic density for N

t has its usual meaning as in 2.1 above

N_N is not measured directly, but rather as N_N/N_{Au} ratio, then a more practical quantity to consider is relative N coverage θ_N/N_N which is simply the product of equations 2.1 and 2.2.

Thus,

$$\theta_N / N_{Au} = I_N / I_{Au} (T_{Au} \sigma_{Au} L^{Q}{}_{Au} / T_N \sigma_N L^{Q}{}_N) \exp(-t / L_{Au}) / 1 - \exp(-t / L_N)(t) \qquad \textbf{(2.4)}$$

where all the parameters have their usual meanings.

The analyzer transmission function T_{Au} and T_N obtained from the instrument are respectively 0.972 and 1.070.

The EALs (i.e. PEALs and QEALs) L_{Au}, L_N, $L^Q{}_{Au}$ and $L^Q{}_N$ were obtained using the SRD-82 [34]. The values obtained are L_{Au} =3.846nm, L_N = 3.109nm, $L^Q{}_{Au}$ = 3.957nm, $L^Q{}_N$ = 3.271nm for sample TT; and L_{Au} =3.843nm, L_N = 3.107nm, $L^Q{}_{Au}$ = 3.954nm, $L^Q{}_N$ = 3.268nm for sample AA.

"Average practical EAL" (PEAL) is explicitly defined as a parameter to be used for overlayers with approximately exponential attenuation. Since DNA is the only type of overlayer being considered here, PEALs refer to attenuation by the DNA film of electrons originating from the specified element [1], e.g. L_N and L_{Au} represent attenuation of electrons from N and Au, respectively. By definition, QEALs refer to electrons originating in a material and attenuated within the material itself, such as Au or N, P, C, and O electrons in a DNA film.

L_{Au} for electrons from the Au substrate in the DNA film were calculated with the following parameters: experimental kinetic energy for Au photoelectrons = 1394.71eV; asymmetry parameter β for electrons Au $4f_{7/2}$ =1.04; theoretical stoichiometry of the DNA samples; band gap energy E_g = 4.8 eV, film density = 0.893g/cm^3 as suggested in ref.1. All these values were used in the SRD-82 software for the calculation of EALs. The total photoelectric cross section σ_{Au} and σ_N (which are respectively 9.58 and 1.80 in units of 13,600 barns) were obtained from tabulated Scofield coefficients [35].

The relative N coverage is proportional to the absolute coverage and so can be used to quantitatively compare coverage. However, the absolute coverage can be determined from θ_N/N_N if the proper film stochiometry is known.

Therefore, absolute coverage:

$$n_{DNA} = \theta_N / N_N (5.892 * 10^{22}\, atoms / cm^3 * 10^{-7}\, cm / nm) / (\#\, Natoms / molecule) \tag{2.5}$$

The terms in brackets ($5.982*10^{22}$atoms/ cm^3 $*10^{-7}$cm/nm) / (# of N atoms/molecule) are collectively called the conversion factor. For example, the conversion factor for sample TT with 40 N atoms per DNA molecule is:

$$(5.982*10^{22}\text{atoms/ cm}^3 *10^{-7}\text{cm/nm}) / 40 \text{ atoms/molecules}$$
$$= 14.73 *10^{13} \text{ molecules/ cm}^2 \text{ nm}$$

The conversion factor obtained using this same expression for sample AA is 5.892 x 10^{13} molecules/ cm^3 nm (considering the fact that there are 100 N atoms per molecule of this ssDNA sample). The units molecules/ cm^2 nm when multiplied by θ_N/N_N which is in nm gives the units of absolute coverage as molecules/ cm^2.

Using equations 2.1 – 2.4 in that order, we obtained the film thickness and relative atomic density for samples TT and AA as shown in Table 2.4. However, the DNA surface coverage cannot be determined with confidence, since the N1s and Au4f signal intensities for both samples are not as reliable, given the limited data set used here. Thus, this parameter can only be more reliably determined after carrying out more controlled experiments with many replicate samples.

Sample	**Film thickness t (nm)**	**Relative atomic density**
TT	**5.022**	**0.0125**
AA	**5.018**	**0.0143**

Table 2.4: Results of the film parameters t and N_N/ N_{Au} as obtained using XPS results and the SRD-82 software. The film thickness of TT being higher than that of AA as explained previously, may be because TT coadsorbed with certain contaminants on the surface which are responsible for the higher film thickness. The fact that relative atomic density of AA is higher than that of TT is evident from the N1s signal intensities of Table 2.3.

FTIR Results

FTIR spectra are shown in Figures 2.6 and 2.7 for samples TT and AA respectively. The results indicate that TT is in fact adsorbed on the gold surface, going by the prominent absorbance 1714 cm^{-1} peak which is specific to the carbonyl groups in thymine [1-2]. This means that the 1714 cm^{-1} feature can be attributed to thymine bases that are contained in the DNA monolayer which are not chemisorbed on the surface [1-2, 24-25]. The very small peak around 1590 cm^{-1} is attributed to chemisorbed thymine [2, 25], indicating that only a very small fraction of the DNA is chemisorbed on the surface. We ascribe the peak at 1474 cm^{-1} to symmetric CH_2 bending mode [36]. Peaks at 1279 cm^{-1} and 1101 cm^{-1} are attributed to asymmetric and symmetric PO_2^- stretch respectively [2].

As for sample AA, the peaks at 1606cm^{-1} and 1651 cm^{-1} are attributed to chemisorbed adenine on gold, in agreement with previous reports on unmodified adenine on gold [23-24]. This supports the above argument from XPS results that sample AA is chemisorbed on the gold surface compared to sample TT. Thus, the FTIR data give qualitative support to the fact that sample TT is adsorbed through the thiol group, while AA is chemisorbed on the gold surface. The peak at 1469 cm^{-1} is ascribed to symmetric CH_2 bending mode, while the one at 1376 cm^{-1} corresponds to symmetric CH_3 bending mode [36]. Peaks at 1257 cm^{-1} and 1105 cm^{-1} are respectively ascribed to asymmetric and symmetric PO_2^- stretch [2].

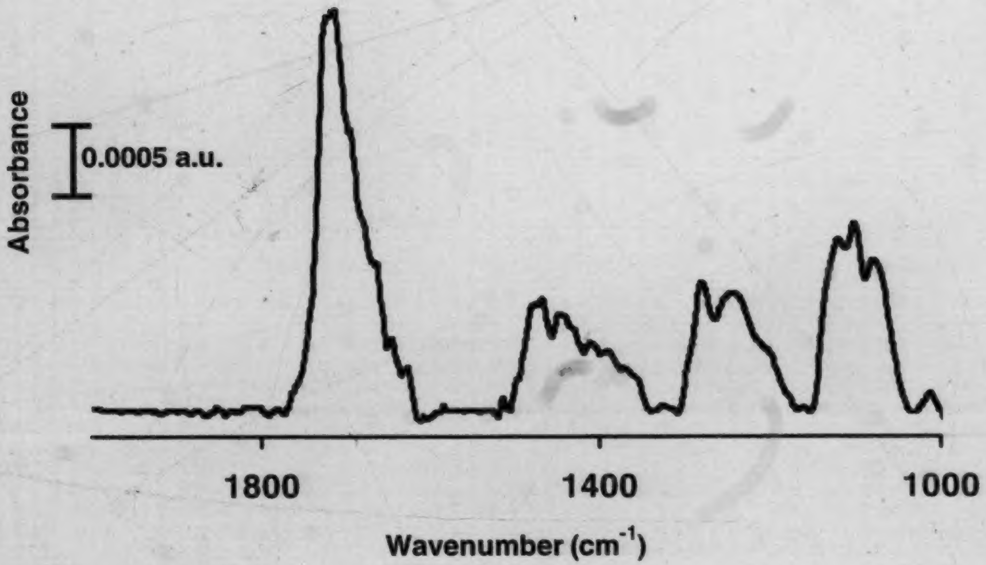


Figure 2.6: FTIR spectrum of 1μM TT in 1TE-NaCl buffer .The peak at 1714 cm^{-1} corresponds to carbonyl groups in free thymine rings. The small peak around 1590 cm^{-1} is attributed to chemisorbed thymine. The peak at 1474 cm^{-1} is attributed to symmetric CH_2 bending mode. Peaks at 1279 cm^{-1} and 1101 cm^{-1} correspond to asymmetric and symmetric PO_2^- stretch, respectively.

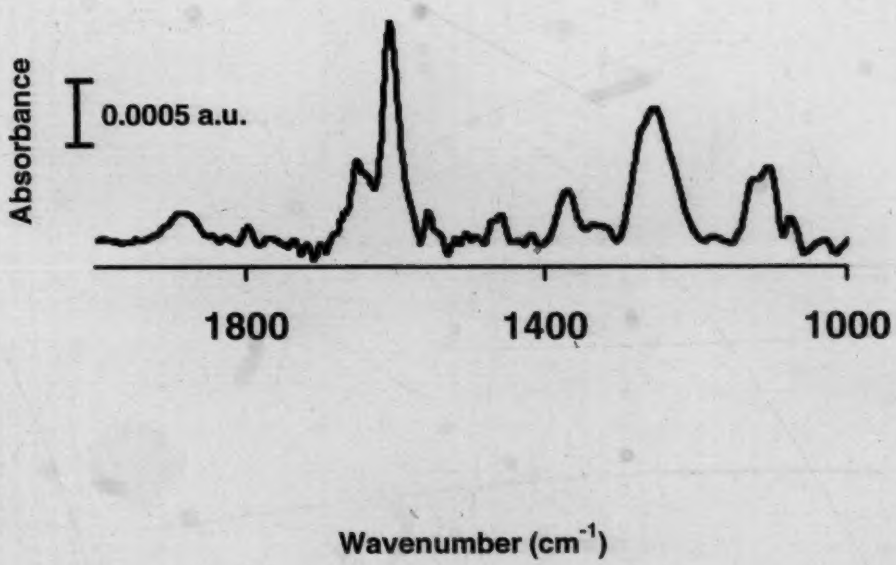


Figure 2.7: FTIR spectrum of 1μM AA in 1TE-NaCl buffer. The carbonyl peak characteristic of TT is conspicuously absent. The peaks at 1606cm^{-1} and 1651 cm^{-1} correspond to NH_2 scissoring and C-N in-plane vibration typical of adenine. The peak at 1469 cm^{-1} corresponds to symmetric CH_2 bending mode, while the one at 1376 cm^{-1} is ascribed to symmetric CH_3 bending mode. Peaks at 1257 cm^{-1} and 1105 cm^{-1} are ascribed to asymmetric and symmetric PO_2^- stretch, respectively.

Conclusions

We have shown how ssDNA can be immobilized on gold surface. Though our results need to be further validated with more experiments, the work presented here is a demonstration of immobilization and characterization of ssDNA on gold. After carrying out the immobilization procedures, we used XPS and FTIR surface analysis methods to characterize the immobilized molecules on gold surface. The XPS results provide some evidence on the adsorption of the molecules on the surface. Our data also showed that the S2p signal for sample TT is half as high as AA. This might mean that sample TT was adsorbed on the surface via the thiol group and so the S atom was buried under the overlying DNA molecules, which resulted in the observed weak S2p signal. In addition, the fact that TT might have coadsorbed with certain contaminants on the gold surface, based on the observed N1s and Au4f signal intensities, could make the surface inaccessible to MCH and sufficiently buries the S atom, further resulting in a very weak S2p signal. On the other hand, AA is believed to be non-specifically chemisorbed on the surface, and so the S atom may not be buried as in TT. In addition, the MCH treatment of AA might have resulted in the displacement of some AA molecules off the surface such that the S2p signal remained essentially the same after MCH treatment. Quantitative XPS analysis (based on the N1s and Au4f signals) using the overlayer model and SRD-82 software, gave film thickness and relative atomic densities for both samples. We found that TT is a relatively thicker film than AA, which may be due to coadsorption of contaminants with TT on the surface, while AA has a higher relative atomic density due to its higher N1s signal. The DNA surface coverage could not be accurately estimated as a result of the suspected contaminant coadsorption, based on the inconsistencies in the observed N1s and Au4f signal intensities. Thus, the results of these experiments have to be validated with replicate analyses of similar samples, as only one sample was analyzed in each case here.

The results from FTIR analysis also give qualitative evidence as to the adsorption of both TT and AA. Wavenumbers characteristic of adsorbed ssDNA samples were identified for both samples.

References

1. D. Y. Petrovykh, H. Kimura-Suda, M. J. Tarlov and L. J. Whitman, *Langmuir,* **20,** 429 (2004).
2. D. Y. Petrovykh, H. Kimura-Suda, L. J. Whitman and M. J. Tarlov, *J. Am. Chem. Soc.* **125,** 5219 (2003).
3. T. M. Herne and M. J. Tarlov, *J. Am. Chem. Soc.* **119,** 8916 (1997).
4. M. J. Tarlov and A. B. Steel, *Biomolecular Films: Design, Function, and Applications;* J. F. Rusling, Ed.; Mercel Dekker: New York, 2003; **Vol. 111.**
5. A. W. Peterson, L. K. Wolf and R. M. Georgiadis, *J. Am. Chem. Soc.* **124,** 14601 (2002).
6. A. B. Steel, R. L. Levicky, T. M. Herne and M. J. Tarlov, *Biophys. J.* **79,** 975 (2000).
7. A. B. Steel, T. M. Herne and M. J. Tarlov, *Anal. Chem.* **70,** 4670 (1998).
8. A. B. Steel, T. M. Herne and M. J. Tarlov, *Bioconj. Chem.* **10,** 419 (1999).
9. J. Wang, G. Rivas, M. Jiang and X. Zhang, *Langmuir,* 15, 6541 (1999).
10. U. Rant, et al. *Langmuir,* **20,** 10086 (2004).
11. R. Levicky, T.M. Herne, M. J. Tarlov and S. K. Satija, *J. Am. Chem. Soc.* **120,** 9787 (1998).
12. A. W. Peterson, R. J. Heaton and R. M. Georgiadis, *Nucl. Acid. Res.* **29,** 5163 (2001).
13. K. A. Peterlinz, R. M. Goergiadis, T. M. Herne and M. J. Tarlov, *J. Am. Chem. Soc.* **119,** 3401 (1997).
14. H. J. Lee, T. T. Goodrich and R. M. Corn, *Anal. Chem.* **73,** 5525 (2001).
15. M. B. Esch, L. E. Locascio, M. J. Tarlov and R. A. Durst, *Anal. Chem.* **73,** 2952 (2001).
16. L. K. Wolf, Y. Gao, and R. M. Georgiadis, *Langmuir,* **20,** 3357 (2004).
17. E. A. Smith, et al. *Langmuir,* **17,** 2502 (2001).
18. S. G. Ray, H. Cohen, R. Naaman and Y. Rabin, *J. Am. Chem. Soc.* **127,** 17138 (2005).
19. T. Aqua, R. Naaman and S. S. Daube, *Langmuir,* **19,** 10573 (2003).

20. C. J. May, H. E. Canavan and D. G. Castner, *Anal. Chem.* **76,** 1114 (2004).
21. C. Y. Lee, H. E. Canavan, L. J. Gamble and D. G. Castner, *Langmuir,* **21,** 5134 (2005).
22. A. V. Saprigin, et al. *Surf. Int. Anal.* **36,** 24 (2004).
23. L. M. Demers, et al. *J. Am. Chem. Soc.* **124,** 11248 (2002).
24. H. Kimura-Suda, D. Y. Petrovykh, M. J. Tarlov and L. J. Whitman, *J. Am. Chem. Soc.* **125,** 9014 (2003).
25. D. Y. Petrovykh, et al. *J. Am. Chem. Soc.* **128,** 2 (2006).
26. Y. D. Zhao, et al. *Anal. Chim. Acta,* **388,** 93, (1999).
27. *http://www.kratos.com/Axis/AXIS_165.html*
28. S. Rapino and F. Zerbetto, *Langmuir,* **21,** 2512 (2005).
29. Private communications with first author of reference 1 above.
30. P. E. Laibinis, et al. *J. Am. Chem. Soc.* **113,** 7152 (1991).
31. D. G. Castner, K. Hinds and D. W. Grainger, *Langmuir,* **12,** 5083 (1996).
32. F. Bansebaa, et al. *Surf. Sci.* **405,** L472 (1998).
33. H. Rieley, et al. *Langmuir,* **14,** 5147 (1998).
34. C. J. Powell, A. Jablonski, *NIST Electron Effective Attenuation Length Database, Version 1.0 (SRD-82);* U.S. Department of Commerce, National Institute of Standards and Technology, Gaithersburg, MD, 2001. This software was gracefully provided free of charge.
35. J. H. Scofield, *J. Electron Spectrosc. Relat. Phenom.* **8,** 129 (1976).
36. R. M. Silverstein, G. C. Bassler and T. C. Morill, *Spectrometric Identification of Organic Compounds,* Fifth edition; John Wiley & Sons, Inc. New York, 1991.

CHAPTER 3

Exploration of the Performance of Selected DFT Functionals for Predicting Vibrational Frequencies of 1, 12-Diphenyl-1, 3, 5, 7, 9, 11-dodecahexayne

Motivation

The motivation for the computational work presented in this thesis is based on the fact that we are interested in modeling the mechanical and electronic properties of polyynes. These compounds are believed to have unique electronic and optical properties. Hence, they can be used as components of molecular electronic devices. So, this work is meant to gain a good understanding of the mechanical and electronic properties of these molecules on the atomic scale. This we believe will enhance our capabilities to design molecular electronic devices of desirable properties.

Introduction

Density Functional Theory (DFT) [1-3] has been gaining popularity over the years as a computationally cost-effective general procedure for studying molecular properties [4-9]. It has been accepted by the traditional *ab initio* quantum chemistry community and efforts have been devoted to refine the methodology and refine the limits of its applicability [5-6, 8, 10-20].

Although a*b initio* quantum mechanical calculations of molecular structure and force fields have contributed a great deal to our understanding of the vibrational spectra of organic molecules [8], most of these calculations are carried out at the Hartree-Fock (HF) level, using the self-consistent field (SCF) approach, with basis sets of double zeta (one STO for inner shell orbital and two STOs for each valence orbital) or double zeta plus polarization (double zeta plus d- or f- type functions added to account for polarization of electron cloud during molecular formation) [4, 8, 11], and the results are marred with systematic errors due to neglect of electron correlation [4, 6-9, 11] and basis set incompleteness [4, 7-8, 11]. These errors are typically removed by use of suitable scaling factors [4, 6-8, 11]. Scaling factors are numbers (usually between 0.85 and 1), used to improve agreement between calculated and experimentally observed vibrational frequencies [3]. The scaling factors also take into account the anharmonicity of the observed vibrational frequencies, thereby yielding results that could be considered as effective force constants [4]. The most popular scaling procedure is the Scaled Quantum Mechanical (SQM) force field method [3-4, 8, 11] introduced by Pulay and co-workers [21], where the force-constant matrix elements, rather than the vibrational frequencies are scaled [3]. The HF-based SQM approach has been successfully applied to study the vibrational spectra of many closed shell organic molecules [4, 8, 11] at their equilibrium geometries [8]. This is because for these systems, HF theory gives qualitatively correct descriptions [4, 8]. However, for many open shell and some conjugated closed shell molecules, these errors are less systematic, and so the SQM approach has been less successful [4, 8, 11].

When formulated in an analogous way to the HF theory, the self-consistent Kohn-Sham (KS) procedure of DFT has a similar size-dependence of cost as the HF theory. But unlike the HF theory, it satisfactorily recovers electron correlation [4, 7-8, 10] in the self-consistent KS procedure through the functionals of the electron density. It also gives good description for systems requiring sophisticated treatment of electron correlation in a conventional a*b initio* approach [4, 8]. Thus, it predicts molecular structural and spectral features in good agreement with experimental results [7, 10]. Results of DFT vibrational frequency calculations can be compared directly with experimental data without any empirical adjustment, in similar or better agreement with the more computationally expensive ab initio methods such as second-order Møller-Plesset perturbation theory (MP2) [10].

In the work presented in this chapter, we compared the performance of selected DFT functionals (fifteen of them in total), at predicting the vibrational frequencies of a model molecule 1, 12-Diphenyl-1, 3, 5, 7, 9, 11-dodecahexayne (which from now on will be referred to as DP-C12). The choice of this model molecule is dictated by our primary goal in this computation work – to model the mechanical and electrical properties of polyyne molecules (this is the subject of chapter 4). Thus, we searched for a molecule with available experimental data and which is simple in structure to be amenable to reasonable computational cost. The molecule that meets these criteria is DP-C12, which was synthesized and characterized by Tykwinski's group, at the University of Alberta, and its experimental data were made available to us [22]. In addition, since we are interested in determining the total energy required to bend a molecule, then the appropriate quantity, which can be used to predict the performance of DFT functionals, is vibrational frequency.

To start with, we tested our method which involves BLYP, B3LYP and BPW91 functionals and a large basis set (6-31G**) on ethylene molecule (which has been the subject of similar calculations) to see what improvement can be achieved in terms of accuracy with the large basis set, compared to the medium basis set (6-31G*) that has always been used in the literature for similar calculations [4, 6-9, 11, 14, 23]. Seeing the improvements, we then proceeded to DP-C12 to test the performance of the selected functionals at predicting its vibrational frequencies. The set of functionals that was used for the DP-C12 includes local, gradient-corrected and hybrid functionals.

The local functionals are: **SVWN,** a combination of Slater exchange functional, also referred to as Local Spin Density (LSD) exchange [1-2, 24-25] and Vosko, Wilk, and Nusair 1980 correlation functional (III) fitting the RPA solution to the uniform electron gas, often referred to as Local Spin Density (LSD) correlation [24, 26], and **SPL,** which combines the local exchange functional of Slater [24-25] and the local (non-gradient corrected) functional of Perdew (1981) [24, 27].

The gradient-corrected functionals are **BP86,** combining Becke's 1988 functional, which includes the Slater exchange along with corrections involving the gradient of the density [24, 28] and the gradient corrections of Perdew, along with his 1981 local correlation functional [24, 29]; **BLYP,** a combination of Becke's 1988 exchange functional, [24, 28] and the correlation functional of Lee, Yang, and Parr which includes both local and non-local terms [24, 30-31] and **BPW91,** which combines Becke's exchange functional, [24, 28] and Perdew and Wang's 1991 gradient-corrected correlation functional [24, 32-34].

The hybrid functionals used (which include a mixture of Hartree-Fock exchange with DFT exchange-correlation), include the Becke's Three Parameter Hybrid Functionals: **B3LYP,** which uses the non-local correlation provided by the LYP expression, and VWN functional (III) for local correlation [24]; **B3P86,** which specifies the same functional with the non-local correlation provided by Perdew 86 [24]; **B3PW91,** which specifies this functional with the non-local correlation provided by Perdew/Wang 91 [24]; and the Becke's One Parameter Hybrid Functionals: **B1B95,** used to specify Becke's one-parameter hybrid functional as defined in the original paper [24, 35]; **B1LYP,** where the LYP correlation functional is used (as described for B3LYP above) [24, 36]; and **MPW1PW91,** which uses modified Perdew-Wang exchange and Perdew-Wang 91 correlation [24, 37].

Other hybrid functionals used are the 1997 hybrid functional of Perdew, Burke and Ernzerhof **PBE1PBE,** which uses 25% exchange and 75% correlation weighting [24, 38]; **B98,** [24, 39-40] which is Becke's 1998 revisions to B97; **B971** - Handy, Tozer and coworkers modification to B97 [24, 41] and **B972** which is Wilson, Bradley and Tozer's modification to B97 [24, 42].

We tested the performance of all these commonly used functionals and chose the one with the best agreement to experimental values. This functional was then used in our modeling calculations for predicting the mechanical and electrical properties of polyynes.

Calculation method

All calculations were performed with the GAUSSIAN 03 suite of programs [43] at the Academic Information and Communication Technology (AICT, formerly CNS) [44], of the University of Alberta. The machines used for the calculations were IBM Model p630 with four processors (Kea) and IBM Model p650 with six processors (Tui). Fifteen different functionals (**SVWN, SPL, BP86, BLYP, BPW91, B3LYP, B3P86, B3PW91, B1B95, B1LYP, MPW1PW91, PBE1PBE, B98, B971 and B972**) were used in this study, as implemented in Gaussian 03.

The large basis set, 6-31G**, was used with all the functionals except for **B3LYP, B3P86, BP86** and **B972,** where a medium basis set 6-31G* was used to avoid expensive computations and the inability of these functionals to find a local energy minimum. In these cases, computations at 6-31G** level converged geometrically, but to a stationary point as indicated by negative frequencies in the results. So, these functionals were used with 6-31G* basis set, with the level of computation increased gradually from STO-3G to 6-31G* to achieve convergence to a local minimum. Though, multiple-bond systems have been reported to converge poorly [45], but why is it that these four functionals did not converge at 6-31G** level of theory while other functionals did? A closer look at these four functionals indicates that their failure to find an energy minimum is perhaps due to the way these functionals were formulated. The hybrid functionals **B3LYP, B3P86** and **B972** have their exchange and correlation components made of local and non-local DFT terms. Since local terms assume constant electron density throughout the molecule, it does not properly account for the inhomogeneity inherent in molecular systems and this is the major drawback of the local density approximation (LDA). So this rather inaccurate assumption may lead to convergence problems with these functionals, more so with such a large basis set. Thus, the molecular system under study is not properly described by these functionals. The other functional **BP86** which is a gradient-corrected functional also has its correlation part substantially made of local terms and so it also fails to converge for the same reason just cited. Hence, the inclusion of substantial amount of local terms in the functionals on the one hand, and the fact that the polyyne system under study is a multiple-bond system which may converge poorly on the other hand, may be responsible for failure of these functionals in finding a real minimum for the molecule. All other functionals that were used with large basis set converged because they lack these deficiencies. This argument becomes more apparent in the section on vibrational frequency of DP-C12 where the performance of all the functionals is analyzed.

Both basis sets were used for the BLYP functionals, as they were chosen as the reference functionals based on its reported excellent performance in the literature [4-9, 11, 15] and our own preliminary results on vibrational frequencies of ethylene. By performance, we mean the accuracy of the computed vibrational frequencies with respect to experimental values. All calculations were carried out with tight convergence criteria. In addition, the quadratically convergent self-consistent field (QC-SCF) approach [24] was used for all calculations to ensure the optimizations find energy minima in all cases.

For those calculations carried out with large basis set, ultrafine numerical integration pruned grids (99 radial and 590 angular points) were used, while those carried out at medium basis set used the default fine numerical integration pruned grids (75 radial and 302 angular points). In all cases, we performed full geometry optimization on the molecule, computed the vibrational frequencies and compared our results with experimental and reported literature values. Molecules were first constructed with HyperChem software [46] (with visualization capability), minimally optimized and the resulting coordinates were ported to GAUSSIAN 03 to prepare the input files for the jobs. The input files were prepared according to GAUSSIAN 03 specifications. Portable batch system (PBS) scripts were written to specify the type of calculations to be performed, the basis set, convergence criteria and what type of results to spell out at the end of the calculations. The scripts were then submitted to the appropriate machine, and after the calculations, the output results are copied into Excel spreadsheet for analysis.

We used the mean percentage absolute deviation (which from now on is referred to as MPAD) as a measure of accuracy.

It was determined by using the expression:

$$(((|V_c - E_v|)/E_v)*100\%)/\#_{vib}$$

where V_c, E_v and $\#_{vib}$ are respectively calculated value, experimental value and number of vibrational modes considered.

Results and Discussion

Vibrational Frequencies of Ethylene

Work has been done and reported on computed vibrational frequencies of ethylene [4, 15]. For simplicity, we therefore first tested our approach on this molecule and compared our results with literature and experiment. It must be pointed out here that, in principle, one should compare calculated harmonic frequencies with 'experimental" harmonic frequencies or the calculated anharmonic frequencies with the observed fundamentals [4, 6, 11, 15]. That is to say, observed vibrational frequencies have anharmonicity incorporated in them, while calculated values are based on pure harmonic approximations. Hence, it is better to include anharmonic corrections in the calculated values where possible, in order to directly compare calculated values with experimentally observed values. In the alternative, one can calculate purely harmonic vibrational frequencies from experimental values where possible, and then compare these harmonic frequencies with calculated values without anharmonicity correction. However, neither the calculation of anharmonic vibrational energy levels, nor reliable derivation of harmonic frequencies from the experimentally observed fundamentals is routinely feasible for polyatomic molecules [4, 6, 11]. Calculation of purely harmonic frequencies from experimental values is only feasible for small molecules like ethylene, and for molecules double the size of ethylene (e. g. octatetraene), calculation of purely harmonic frequencies from experimental values is rather difficult. Therefore, only the results obtained for ethylene have the experimental harmonic frequencies calculated. Generally speaking, **BLYP** is known to give better agreement between calculated and experimentally observed vibrational frequencies, even without the use of scaling factors [4, 6-9, 11, 15, 23]. As stated by Zhou et al. in ref. 4, the agreement between **BLYP** and observed vibrational frequencies may be due to error cancellation. **BLYP** predicts bond lengths that are slightly longer than experimental values, so the computed force constants and vibrational frequencies would be slightly lower than what they ought to be. Since the effect of anharmonicity is also to lower the vibrational frequencies, then, it follows that the high level of conformity between

BLYP frequencies and experimental values is due to the overestimation of bond lengths by the **BLYP** method. So, because of the high level of agreement between **BLYP** vibrational frequencies and experimentally observed values, the results of our calculations on DP-C12 are directly compared with experimental values.

Table 3.1 shows the results obtained from previous calculations on ethylene (Figure 3.1) as compared with reported experimentally observed values, in conjunction with peak assignments and the anharmonicity correction factors. Harmonic vibrational frequencies were obtained in this case by using the expression $v = w(1-x)$; where v is the observed fundamental frequency, w is the harmonic frequency and x is the anharmonicity factor which is 0.04, 0.02 and 0.015 for C-H stretching, bending and C-C stretching, respectively [47].

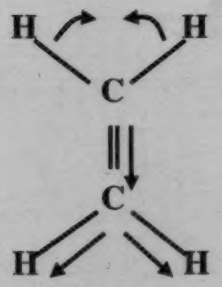


Figure 3.1: Cartoon showing the structure of ethylene molecule and some of the possible vibrational modes. This polyatomic molecule of six atoms should have twelve normal vibrational modes.

Observed [4, 47] (cm^{-1})	**Harmonic [15] (cm^{-1})**	**Assignment [48]**	**Anharmonicity factor**
826	**843**	**CH2 rock**	**0.02**
940	**959**	**CH2 wag**	**0.02**
949	**969**	**CH2 wag**	**0.02**
1023	**1044**	**CH2 twist**	**0.02**
1220	**1245**	**CH2 rock**	**0.02**
1342	**1370**	**CH2 scis**	**0.02**
1444	**1473**	**CH2 scis**	**0.02**
1630	**1655**	**CC str**	**0.015**
3021	**3147**	**CH2 sym-str**	**0.04**
3026	**3153**	**CH2 sym-str**	**0.04**
3086	**3232**	**CH2 asym-str**	**0.04**
3105	**3234**	**CH2 asym-str**	**0.04**

Table 3.1: The results obtained from previous DFT calculations [15] as compared with reported experimentally observed values [4, 47]. Also shown are the peak assignments and the anharmocity correction factors.

The results of our own calculations are presented below. Only three functionals B3LYP, BLYP and BPW91 were used here, with the "spectroscopic" basis set (6-31G*). This basis set is termed "spectroscopic" by some authors probably because it has been in common use for computing the vibrational frequencies of organic molecules since its results are superior to those of smaller basis sets and further enlarging the basis set increases the computational cost [4, 6, 8-9, 23]. The first two of these three functionals have been used in the literature. The results obtained were compared with experimental and reported literature values and the deviations of our values are given in both cases. It must be pointed out that literature references, for the most part, used the mean absolute deviation of non-CH stretching modes as their measure of accuracy [4, 6-7, 9, 11, 23], since larger deviations were found for the C-H stretching modes because these are strongly affected by anharmonicity. In our work, we used MPAD, which cuts across all modes. Thus, our reports of deviations are more representative than literature values of mean absolute deviation of non-CH stretching modes.

Experimentally observed values [4, 47] (cm^{-1})	B3LYP (cm^{-1})	Theoretical results from Ref. 4 (cm^{-1})	PAD (%) relative to experimental values in column 1	PAD (%) relative to purely harmonic frequencies derived from experiment [15]
826	835.242	835	1.119	0.920
940	956.407	956	1.745	0.270
949	976.361	976	2.883	0.760
1023	1070.323	1070	4.626	2.521
1220	1248.14	1248	2.307	0.252
1342	1396	1396	4.024	1.898
1444	1494.385	1495	3.489	1.452
1630	1720.648	1720	5.561	3.967
3021	3151.239	3151	4.311	0.135
3026	3166.71	3167	4.650	0.435
3086	3221.544	3221	4.392	0.324
3105	3247.119	3247	4.577	0.406
MPAD (%)			3.640	1.112

Table 3.2: The results obtained from our calculations (with **B3LYP/6-31G***) as compared to reported experimental and theoretical literature values. Deviations of our results from experimentally observed and harmonic vibrational frequencies are also shown. PAD is the percentage absolute deviation of the individual vibrational frequencies.

Experimentally observed values [4, 47] (cm^{-1})	BLYP (cm^{-1})	Theoretical results from Ref. 15 (cm^{-1})	PAD (%) relative to experimental values in column 1	PAD (%) relative to purely harmonic frequencies derived from experiment [15]
826	814.683	814	1.370	3.359
940	907.347	907	3.474	5.386
949	940.356	940	0.911	2.956
1023	1040.189	1040	1.680	0.365
1220	1216.857	1217	0.258	2.260
1342	1359.374	1360	1.295	0.776
1444	1458.474	1458	1.002	0.986
1630	1664.199	1664	2.098	0.556
3021	3068.215	3069	1.563	2.503
3026	3081.586	3082	1.837	2.265
3086	3133.33	3134	1.534	3.053
3105	3160.048	3160	1.773	2.287
MPAD (%)			1.566	2.229

Table 3.3: The results obtained from our calculations (with **BLYP/6-31G***) as compared to reported experimental and theoretical literature values. Deviations of the results from experimentally observed and harmonic vibrational frequencies are also shown for comparison. PAD is the percentage absolute deviation of the individual vibrational frequencies.

Experimentally observed values [4, 47] (cm^{-1})	BPW91 (cm^{-1})	PAD (%) relative to experimental values in column 1	PAD (%) relative to purely harmonic frequencies derived from experiment [15]
826	811.923	1.704	3.686
940	909.895	3.203	5.120
949	939.846	0.965	3.009
1023	1041.731	1.831	0.217
1220	1211.794	0.673	2.667
1342	1361.298	1.438	0.635
1444	1451.31	0.506	1.473
1630	1671.016	2.516	0.968
3021	3085.56	2.137	1.952
3026	3100.967	2.477	1.650
3086	3157.343	2.312	2.310
3105	3183.449	2.527	1.563
MPAD (%)		1.857	2.104

Table 3.4: The results obtained from our calculations (with **BPW91/6-31G***) as compared to reported experimental literature values. Deviations of these results from experimentally observed and harmonic vibrational frequencies are also shown. PAD is the percentage absolute deviation of the individual vibrational frequencies.

From the results above, we found a perfect agreement between our values and reported literature values. It can also be deduced from our results that, on the one hand, B3LYP harmonic frequencies are closer to "experimental" harmonic frequencies than BLYP harmonic frequencies, while on the other hand, BLYP harmonic frequencies are closer to experimentally observed frequencies than B3LYP harmonic frequencies, an observation that has also been made previously [4]. This same trend is observed when 6-31G** basis set was used. But since direct comparison of calculated frequencies and the experimentally observed results is more straightforward and practical, than comparing B3LYP results with the observed results, the direct comparison path is always taken by various researchers, using the BLYP functionals [4, 6-9, 11, 15, 23].

Based on the above results, we then used a large basis set 6-31G** with the same three DFT functionals in an effort to get better agreement with experimental values. The large basis set expectedly gave better results than 6-31G* as indicated by the smaller MPAD values. Thus, the large basis set, though moderately more expensive, gave more accurate results than the medium basis set earlier used. Our results from calculations with the large basis set are presented in the following tables.

Experimentally observed values [4, 47] (cm^{-1})	B3LYP (cm^{-1})	PAD (%) relative to experimental values in column 1 (6-31G**)	PAD (%) relative to experimental values in column 1 (6-31G*)	PAD (%) relative to purely harmonic frequencies derived from experiment [15] (6-31G**)	PAD (%) relative to purely harmonic frequencies derived from experiment [15] (6-31G*)
826	831.667	0.686	1.119	1.344	0.920
940	961.28	2.264	1.745	0.238	0.270
949	976.876	2.937	2.883	0.813	0.760
1023	1069.519	4.547	4.626	2.444	2.521
1220	1241.642	1.774	2.307	0.270	0.252
1342	1388.634	3.475	4.024	1.360	1.898
1444	1483.083	2.707	3.489	0.685	1.452
1630	1715.544	5.248	5.561	3.658	3.967
3021	3144.776	4.097	4.311	0.071	0.135
3026	3160.533	4.446	4.650	0.239	0.435
3086	3220.542	4.360	4.392	0.355	0.324
3105	3246.473	4.556	4.577	0.386	0.406
MPAD (%)		3.425	3.640	0.988	1.112

Table 3.5: The results obtained from our calculations **(B3LYP/6-31G**)** as compared with reported experimental and theoretical literature values. Deviations of our results from experimentally observed and harmonic vibrational frequencies are shown. PADs from previously used medium basis set are also shown for comparison. PAD is the percentage absolute deviation of the individual vibrational frequencies. It must be noted that, contrary to expectation, some of the C-H stretch vibrational frequencies have larger deviations at 6-31G** level than 6-31G* level.

Experimentally observed values [4, 47] (cm^{-1})	BLYP (cm^{-1})	PAD (%) relative to experimental values in column 1 (6-31G**)	PAD (%) relative to experimental values in column 1 (6-31G*)	PAD (%) relative to purely harmonic frequencies derived from experiment [15] (6-31G**)	PAD (%) relative to purely harmonic frequencies derived from experiment [15] (6-31G*)
826	811.333	1.776	1.370	3.756	3.359
940	913.541	2.815	3.474	4.740	5.386
949	942.132	0.724	0.911	2.773	2.956
1023	1040.297	1.691	1.680	0.355	0.365
1220	1210.334	0.792	0.258	2.784	2.260
1342	1352.725	0.799	1.295	1.261	0.776
1444	1447.843	0.266	1.002	1.708	0.986
1630	1658.935	1.775	2.098	0.238	0.556
3021	3065.415	1.470	1.563	2.592	2.503
3026	3079.052	1.753	1.837	2.345	2.265
3086	3136.304	1.630	1.534	2.961	3.053
3105	3163.355	1.879	1.773	2.184	2.287
MPAD (%)		1.448	1.566	2.308	2.229

Table 3.6: The results obtained from our calculations **(BLYP/6-31G**)** as compared with reported experimental and theoretical literature values. Deviations of our results from experimentally observed and harmonic vibrational frequencies are shown. PADs from previously used medium basis set are also shown for comparison. PAD is the percentage absolute deviation of the individual vibrational frequencies. Here too, some vibrational frequencies (especially C-H stretch) have arger deviations at 6-31G** level than 6-31G* level.

Experimentally observed values [4, 47] (cm^{-1})	**BPW91 (cm^{-1})**	**PAD (%) relative to experimental values in column 1 (6-31G**)**	**PAD (%) relative to experimental values in column 1 (6-31G*)**	**PAD (%) relative to purely harmonic frequencies derived from experiment [15] (6-31G**)**	**PAD (%) relative to purely harmonic frequencies derived from experiment [15] (6-31G*)**
826	**807.718**	**2.213**	**1.704**	**4.185**	**3.686**
940	**914.822**	**2.679**	**3.203**	**4.607**	**5.120**
949	**940.496**	**0.896**	**0.965**	**2.942**	**3.009**
1023	**1041.133**	**1.773**	**1.831**	**0.275**	**0.217**
1220	**1203.953**	**1.315**	**0.673**	**3.297**	**2.667**
1342	**1353.203**	**0.835**	**1.438**	**1.226**	**0.635**
1444	**1438.811**	**0.359**	**0.506**	**2.321**	**1.473**
1630	**1665.098**	**2.153**	**2.516**	**0.610**	**0.968**
3021	**3083.138**	**2.057**	**2.137**	**2.029**	**1.952**
3026	**3098.659**	**2.401**	**2.477**	**1.723**	**1.650**
3086	**3160.164**	**2.403**	**2.312**	**2.223**	**2.310**
3105	**3186.459**	**2.623**	**2.527**	**1.470**	**1.563**
MPAD (%)		**1.809**	**1.857**	**2.242**	**2.104**

Table 3.7: The results obtained from our calculations **(BPW91/6-31G**)** as compared with reported experimental and theoretical literature values. Deviations of our results from experimentally observed and harmonic vibrational frequencies are shown. All vibrational frequencies are in units of cm^{-1}. PADs from previously used medium basis set are also shown for comparison. PAD is the percentage absolute deviation of the individual vibrational frequencies. Some vibrational frequencies (especially C-H stretch) have larger deviations at 6-31G** level than 6-31G* level.

Summing it all up, we have the summary of the performances of the three functionals as shown in Table 3.8.

Method	**Computation time (x 10^3s)**	**MPAD (%)**
BLYP/6-31G**	**5**	**1.448**
BLYP/6-31G*	**4**	**1.566**
BPW91/6-31G**	**5**	**1.809**
BPW91/6-31G*	**4**	**1.857**
B3LYP/6-31G**	**5**	**3.425**
B3LYP/6-31G*	**4**	**3.640**

Table 3.8: This table shows the overall deviations of our values from experimentally observed values as obtained for each of the methods. BLYP/6-31G** has the best overall performance, followed by BLYP/6-31G*. Thus, by using BLYP with 6-31G** instead of 6-31G*, there can be an improvement in accuracy of about 8%. One interesting trend observed here is that BLYP/6-31G* even performed better than the other two functionals with a large basis set. In addition, it can be concluded that expectedly, when large basis set is used with these functionals, a better agreement can be obtained with experimental values, compared to the case of medium basis set.

A closer look at the performance of the different functionals and the basis sets, indicates that large basis set gives better agreement between calculated (harmonic) and experimentally observed (harmonic and anharmonic) vibrational frequencies. In addition, it can be deduced from the above results that the gradient-corrected functionals (**BLYP** and **BPW91**) outperformed the hybrid functionals **B3LYP**. The differences in performance of the functionals may be rationalized based on the way they were formulated. The **B3LYP** performed the least because its exchange part is a mixture of the exact HF, LDA and GGA terms, while its correlation part is made of a good fraction of local terms which, as stated previously, has a major shortcoming in describing molecular systems by assuming a constant electron density as a function of position. Since HF is inaccurate for the correct qualitative description of some conjugated organic molecules [4], and the LDA term in the correlation part is also an inadequate descriptor of the electron density of molecular systems, then this functional would give less accurate vibrational frequencies for such systems. **BLYP** and **BPW91** on the other hand are both gradient-corrected functionals [24], whose correlation components basically have no local terms. Hence, these can give more correct description for the molecular system under study and more accurate vibrational frequencies.

Generally, in terms of computation time, the calculations carried out with a large basis set are about a thousand seconds more expensive per run than those at medium basis set, as seen in the above table.

Using the above case as a benchmark, we then used the functionals with the best overall agreement with experimental values **(BLYP/6-31G**)** as reference functionals to study our model system.

Vibrational frequencies of DP-C12

We computed vibrational frequencies of this molecule and compared results with available experimental data [22]. Fifteen different DFT functionals (**SVWN, SPL, BP86, BLYP, BPW91, B3LYP, B3P86, B3PW91, B1B95, B1LYP, MPW1PW91, PBE1PBE, B98, B971 and B972**) were tested. Both basis sets were used for the **BLYP** functionals, being the reference functional. Than Luu, of Professor

Tykwinski's group, generously made the experimental data on the molecule available to us. A total of thirteen frequency peaks were used for comparison between our values and experimental values.

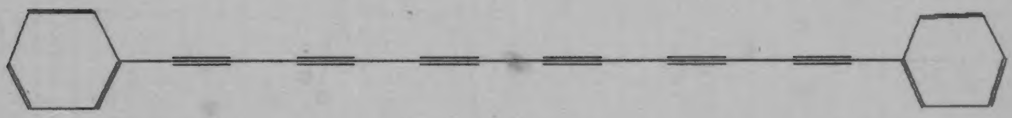


Figure 3.2: The chemical structure of DP-C12 used in this study. The structure is similar to our straight chain polyynes except for the 'flanking' phenyl rings.

The following table shows the deviations in increasing order, of our calculated vibrational frequencies from the experimentally observed values, as well as the computation times.

Functionals	**Basis set**	**MPAD (%)**	**CPU time/ iteration (x10^{3}s)**
BLYP	**6-31G****	**0.338**	**6.3**
BLYP	**6-31G***	**0.431**	**2.6**
BPW91	**6-31G****	**0.478**	**6.7**
BP86	**6-31G***	**0.774**	**2.5**
SPL	**6-31G****	**2.776**	**5.9**
B98	**6-31G****	**2.971**	**6.7**
SVWN	**6-31G****	**3.321**	**4.5**
B971	**6-31G****	**4.364**	**5.4**
B972	**6-31G***	**4.563**	**4.8**
B3LYP	**6-31G***	**4.683**	**3.2**
B3PW91	**6-31G****	**4.858**	**9.4**
B3P86	**6-31G***	**5.038**	**3.3**
B1B95	**6-31G****	**5.482**	**16.0**
B1LYP	**6-31G****	**5.496**	**6.4**
PBE1PBE	**6-31G****	**5.651**	**8.4**
MPW1PW91	**6-31G****	**5.860**	**9.2**

Table 3.9: Results obtained with the fifteen DFT functionals on the model compound (DP-C12). The functionals are arranged in order of decreasing performance based on MPAD values. In accord with previous results, BLYP functional also has the best overall performance with both bases sets. The CPU time per iteration for each functional is also shown. This represents the time taken for the first QC-SCF iteration of the highest level of theory used for DFT geometry optimization in each case.

The performance of the functionals reveals an interesting trend. For the purpose of this exercise, we can broadly group the functionals into three. The groups are: the best DFT functionals, with MPAD values < 1; average DFT functionals, with MPAD values in the range 2 - 4; and the poor DFT functionals, having MPAD values > 4. It can be easily seen that the best functionals, - **BLYP, BPW91** and **BP86** are gradient-corrected functionals, while the three average functionals, are the two local functionals used – **SVWN** and **SPL**, with only one hybrid functional **B98**. All other hybrid functionals used constitute the poor functionals. So basically, gradient-corrected functionals generally outperformed local and hybrid functionals. Thus, it can be reasonably argued that the level of performance of these functionals is related to the way they were formulated.

Since gradient-corrected functionals are based on the generalized-gradient approximation (GGA), they give more accurate description of molecular systems by correcting for the variation of electron density with position in the LDA approach. We can therefore say that the formulation of the gradient-corrected funtionals correctly accounts for the inhomogeneous nature of electrons in molecular systems, hence it gives more accurate vibrational frequencies. The local functionals performed less well because their exchange and correlation terms are based on LDA, which assumes a uniform electron gas model [3], whose density is a constant or a slowly varying function of position. This approximation is rather inaccurate for describing the ideal nature of electrons in molecular systems. Thus, the computed vibrational frequencies are not as accurate. In the same vein, it may be rationalized that the hybrid functionals performed the worst because their exchange components are blends of HF, LDA and GGA terms on the one hand, and their correlation parts consist of substantial fraction of the inaccurate local terms in conjunction with gradient terms on the other hand. Since HF method does not give correct qualitative description for some conjugated organic systems [4], and a local approximation is inaccurate for describing molecular systems, the incorporation of these in the hybrid functionals may be responsible for the less accurate computed vibrational frequencies.

The only exception in the list is **B98**, a hybrid functional whose performance is similar to those of local functionals. This behaviour may also be argued considering the formulation of this functional. Its exchange and correlation parts are composed of 75% or 100% local terms making it look more or less like a local functional. Thus, this functional eventhough is termed a hybrid functional, its formulation is such that the exchange and correlation components are almost completely, if not wholly, local. It is not surprising therefore that this functional performed like the original local functionals.

In terms of computation cost (CPU time per iteration), those functionals used with 6-31G* basis set took the least time (about 50 minutes of CPU time per iteration), except B972 which took about 70 minutes. This is followed by the local functionals which took about 80 minutes. The gradient-corrected functionals took about one hundred minutes per iteration, while most hybrid functionals took over two hours per iteration, except B1LYP functionals, which took about one hundred minutes. Thus, those functionals used with 6-31G* basis set took the least time per iteration, due to the fact that these calculations were carried out at a lower level of theory. The local functionals that use the LDA approximation, which is not so accurate in describing molecular systems, also took relatively short time at the expense of accuracy. The approximation used by the gradient-corrected functionals is more accurate and so these functionals require more time in finding the optimized geometry for the molecule and computing its vibrational frequencies as shown in the previous table. It may also be rationalized that most hybrid functionals took the longest time per iteration because they include both HF, local and gradient terms in their formulations. The computed vibrational frequencies of the best functionals **BLYP** are shown in Table 3.10 as well as the corresponding experimental values. The peak assignments are also displayed with the deviations of calculated values from experimental values.

Experimental [22] (cm^{-1})	Calculated values (BLYP/6-31G**) (cm^{-1})	Peak assignments [49-51]	PAD (%) relative to experimental values of column 1
524.24	519.07	C≡ C- C rock	0.986
680.22	677.75	C≡ C str	0.363
680.22	677.79	C≡ C str	0.357
746.43	745.63	C≡ C- C scis	0.107
909.35	910.15	C≡ C- scis	0.088
997.35	1002.33	C≡ C- C rock	0.499
1025.58	1021.73	C≡ C- C rock	0.376
1175.53	1174.64	C≡ C str	0.076
1441.32	1434.12	C≡ C str	0.499
1441.32	1434.12	C≡ C str	0.499
1485.86	1485.26	C≡ C- scis	0.040
2159.18	2156.50	C≡ C str	0.124
2174.44	2182.70	C≡ C str	0.380
		MPAD (%)	0.338

Table 3.10: The results obtained for DP-C12 from our calculations **(BLYP/6-31G**)** as compared with reported experimental values. Deviations of our results from experimentally observed values are shown for comparison. PAD is the percentage absolute deviation of the individual computed vibrational frequencies from experimentally observed values.

From the experimentally observed wavenumbers, the force constant can be computed using the expression:

$$\nu = 1/2\pi c(k/\mu)^{1/2}$$

where ν is the wavenumber in cm^{-1}

c is the velocity of light in cm/s

k is the force constant in N/m

μ is the reduced mass ($\mu = m_1 m_2 / m_1 + m_2$) in kg

The force constant for the C≡ C-C bending vibrational mode (524.24 cm^{-1}) in the table was found to be 97.15 N/m, while the stretching vibrational mode (2174.44 cm^{-1}) gave a force constant of 1671.43 N/m. As can be seen, the bending force constant is far less than the stretching force constant, indicative of the relative energies required to bring about these molecular vibrations. These values should have some relationship with the overall or effective bending and stretching force constants of polyyne molecules. If the molecules are assumed to behave like unit springs joined together in series (each of C-C, C≡ C, C–H etc. representing a unit spring), then we can estimate the effective force constants of each polyyne molecule by using the expressions given below.

$$K_{eff} = k / n_d \qquad \textbf{(3.1)}$$

K_{eff} is the overall or effective force constant, k is the force constant of each unit (e.g. C≡ C) and n_d is the number of units joined together in series [52]. To a first approximation, we can assume that the force constant of identical units (e.g. terminal C≡ C and internal C≡ C) in the molecules are the same, then the effective force constants of similar units can be estimated from this equation. When the units directly joined together are different, so that they have different force constants, e.g. (C≡ C and C–H), the following expression is used:

$$K_{eff} = (1/k_1 + 1/k_2 + 1/k_3 + \ldots\ldots + 1/k_n)^{-1} \qquad \textbf{(3.2)}$$

where k_1, k_2 and k_n are the unit force constants of the different units joined together in series [53].

The effective bending and stretching force constants for a polyyne molecule can be estimated using the force constants obtained from the experimental vibrational modes (values indicated above), and assuming a C≡ C–H bending mode of 630 cm^{-1} (k = 21.59 N/m), a C–C stretching of 1200 cm^{-1} (k = 509.05 N/m) and a C–H stretching of 3300 cm^{-1} (k = 592.26 N/m) [49]. This can be done by joining the constituents units together (e.g. C≡ C–C and C≡ C–H for bending, and C–C, C≡ C and C–H for stretching) to make each polyyne molecule, then the respective force constants stated above are substituted in the given equations to obtain the effective force constants. This is interesting, realizing that we can derive functions, which have physical significance from this approximation, these can then be used to describe the force constant curves resulting from our computations on polyynes. The estimated effective force constants obtained using the equations above are shown in Table 3.11.

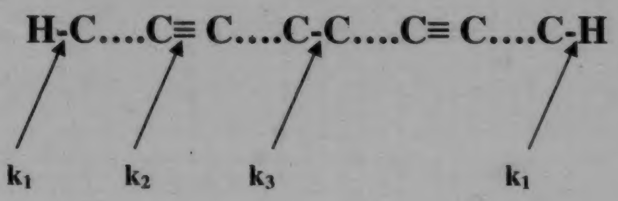


Figure 3.3: This figure shows a typical polyyne molecule being made up of constituent units, each with its characteristic force constant as shown. The overall force constant can be derived from these constituent force constants as stated above.

# of carbon atoms	Bending force constant (N/m)	Stretching force constant (N/m)
4	10.80	152.95
6	9.72	109.88
8	8.84	85.73
10	8.10	70.29
12	7.48	59.56

Table 3.11: The estimated effective bending and stretching force constants for C_4-C_{12} polyyne molecules. The values were obtained by joining together the constituent units that make up each polyyne molecule and calculating the effective force constants from the force constant of each constituent unit.

From the above equations, we derive the following functions (which of course fit the data in Table 3.11 reasonably well). These functions will be used to fit force constant curves in our subsequent calculations.

$$K_{eff} = (2/k_1 + ((n-4)/2)/k_2)^{-1} \qquad (3.3)$$

This function describes bending in polyynes, n is the number of carbon atoms per molecule, and k_1 and k_2 represent force constants for C≡ C–H and C≡ C–C bending, respectively.

$$K_{eff} = (2/k_1 + (n/2)/k_2 + ((n/2)-1)/k_3)^{-1} \qquad \textbf{(3.4)}$$

This function defines stretching in polyynes, with k_1, k_2 and k_3 representing force constants for C–H, C≡ C and C–C stretching, respectively.

Conclusions

We have explored the performance of three commonly used functionals with both medium and large basis sets to compute the vibrational frequencies of ethylene. Our results from the calculations with 6-31G* are in perfect agreement with literature reports. It was revealed that **BLYP/6-31G**** has the best overall performance, for its results are the closest to experimentally observed values. The two gradient-corrected functionals used here performed better than the hybrid functionals. This is based on the differences in the way the two classes of functionals were formulated. The exchange and correlation parts of the gradient-corrected functionals are made of gradient terms, which are more accurate in describing molecular systems. The hybrid functional has its exchange part made of HF, LDA and GGA terms, while its correlation part is composed of both local and non-local terms. The inaccuracy of HF and LDA in describing molecular systems is believed to be responsible for the less accurate vibrational frequencies computed by these functionals. Our results also show that for computing the vibrational frequencies of ethylene, using **BLYP** with 6-31G** instead of 6-31G* basis set, gives about 8% improvement in accuracy, though relatively more computationally expensive.

When the set of selected functionals were used with the large basis set for computing vibrational frequencies of DP-C12, **BLYP/6-31G**** also has the best overall performance, followed by **BLYP/6-31G***. Our results show that the performance of the functionals has correlation with the way they were formulated. The best functionals are the gradient-corrected functionals, followed by the local functionals, while the hybrid functionals performed the least. Since the GGA approximation gives correct qualitative description of molecular systems and correct the inherent inhomogeneity in the LDA approximation, functionals that are so formulated would be able to give accurate vibrational frequencies as observed with the gradient-corrected functionals. The local functionals gave less accurate frequencies since they are based on the inaccurate LDA approximation. The hybrid functionals performed the least as a result of the fact that their exchange components contain HF, LDA and GGA terms, while their correlation parts consist of substantial fraction of local terms, which are rather less accurate. Since the HF method and the LDA are both less accurate for qualitatively describing some molecular systems, which may include DP-C12, then the computed vibrational frequencies would not be as accurate. The hybrid functional **B98** is an exception in the list because its performance is similar to those of local functionals. This is because this functional is formulated in an analogous way to the local functionals – its exchange and correlation parts are composed of 75% or 100% local terms making it look more or less like a local functional. Thus, this functional falls in the same performance rating as the original local functionals.

Considering the computation time (CPU time per iteration), those functionals used with 6-31G* basis set took the least time (about 50 minutes of CPU time per iteration), with the exemption of B972 which took about 70 minutes. The local functionals took about 80 minutes, while the gradient-corrected and most hybrid functionals took about two hours or more. Those functionals used with 6-31G* took the least time, based on the relatively lower level of theory employed, or perhaps due to the fact that these calculations were broken into four steps to ensure convergence to a local minimum. Next in the series are the local functions which took about 80 minutes. This we believe is because these functionals assume less accurate approximations, and so compute the vibrational frequencies faster than most other functionals while sacrificing accuracy. The gradient-corrected functionals use better and more accurate approximation, and so require more time to give more accurate results. The hybrid functionals use mixtures of both approximations in conjunction with HF theory; this caused them also to take such a long time per iteration.

We obtained force constants from the experimentally observed wavenumbers and found the bending force constant to be less than that of stretching, indicating the relative energies needed to cause these vibrations. Also, the effective bending and stretching force constants for different polyyne molecules were estimated from the experimentally observed wavenumbers. We also derived functions which can be used to fit force constant curves and attach physical meaning to the bending and stretching of the polyyne molecules.

This work demonstrates for the first time, the performance of different DFT functionals at predicting vibrational frequencies of a polyyne molecule. The exploration produced **BLYP** as the best overall functional for the molecular system studied. Consequently, **BLYP/6-31G**** will be used for modeling the mechanical and electronic properties of straight chain polyynes. This is the subject of the next chapter.

References

1. P. Hohenberg and W. Kohn, *Phys. Rev.* **B136,** 864 (1964).
2. W. Kohn and L. J. Sham, *Phys. Rev.,* **A140,** 1133 (1965).
3. I. N. Levine, *Quantum Chemistry*, (pg. 573) Prentice Hall (2000).
4. X. Zhou, C. J. M. Wheeless, R. Liu, *Vibrational Spectroscopy,* **12,** 53 (1996).
5. X. Zhou and R. Liu, *Vibrational Spectroscopy,* **12,** 65 (1996).
6. X. Zhou, S. J. Mole, R. Liu, *Vibrational Spectroscopy,* **12,** 73 (1996).
7. J. Higgins, X. Zhou, R. Liu, *Spectrochim Acta Part A,* **53**, 721 (1996).
8. X. Zhou, R. Liu, *Spectrochim Acta Part A,* **53,** 259(1997).
9. S. J. Mole, X. Zhou, J. G. Wardeska and R. Liu, *Spectrochim Acta Part A,* **52,** 1211 (1996).
10. M. D. Halls and H. B. Schlegel, *J. Chem. Phys.* **109,** 10587 (1998).
11. C. J. M. Wheeless, X. Zhou, R. Liu, *J. Chem. Phys.* **99,** 12488 (1995).

12. J. Andzelm and E. Wimmer, *J. Chem. Phys.* 96, 1280 (1996).
13. A. D. Becke, *J. Chem. Phys.* 96, 2155 (1992).
14. B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.* **97,** 7846 (1992).
15. B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.* **98,** 5612 (1993).
16. P. M. W. Gill, B. G. Johnson and J. A. Pople, *Chem. Phys. Lett.* **197,** 449 (1993).
17. T. Ziegler, *Chem. Rev.* **91,** 651 (1991).
18. N. C. Handy, et al. *Chem. Phys. Lett.* **197,** 506 (1992).
19. C. W. Murray, G. J. Laming, N. C. Handy and R. D. Amos, *Chem. Phys. Lett.* **199,** 551 (1992).
20. J. Baker, A. Scheiner and J. Andzelm, *Chem. Phys. Lett.* **216,** 380 (1993).
21. P. Pulay, et al. *J. Am Chem. Soc.* 105, 7037 (1983).
22. T. Luu, et al. Org. Lett. 7, 51 (2005).
23. B. Ball, X. Zhou, R. Liu, Spectrochim Acta Part A, **52**, 1803 (1996).
24. http://www.gaussian.com/g_ur/k_dft.htm
25. J. C. Slater, *Quantum Theory of Molecular and Solids. Vol. 4: The Self-Consistent Field for Molecular and Solids* (McGraw-Hill, New York, 1974).
26. S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys* . **58**, 1200 (1980).
27. J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
28. A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
29. J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
30. C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
31. B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.* **157**, 200 (1989).
32. K. Burke, J. P. Perdew, and Y. Wang, in *Electronic Density Functional Theory: Recent Progress and New Directions*, Ed. J. F. Dobson, G. Vignale, and M. P. Das (Plenum, 1998).
33. J. P. Perdew, in *Electronic Structure of Solids* ' 91, Ed. P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991) 11.
34. J. P. Perdew, K. Burke, and Y. Wang, *Phys. Rev. B* **54**, 16533 (1996).
35. A. D. Becke, *J. Chem. Phys.* **104**, 1040 (1996).

36. C. Adamo and V. Barone, *Chem. Phys. Lett.* **274**, 242 (1997).
37. C. Adamo and V. Barone, *J. Chem. Phys.* **108**, 664 (1998).
38. J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
39. A. D. Becke, *J. Chem. Phys.* **107**, 8554 (1997).
40. H. L. Schmider and A. D. Becke, *J. Chem. Phys.* **108**, 9624 (1998).
41. F. A. Hamprecht, A. J. Cohen, D. J. Tozer, and N. C. Handy, *J. Chem. Phys.* **109**, 6264 (1998).
42. P. J. Wilson, T. J. Bradley, and D. J. Tozer, *J. Chem. Phys.* **115**, 9233 (2001).
43. M. J. Frisch, et al. *Gaussian 03, Revision B.04,* Gaussian, Inc., Pittsburgh PA, 2003.
44. *http://www.ualberta.ca/CNS/*
45. *HyperChem Computational Chemistry, Practical Guide, Theory and Methods,* Hypercube, Inc., 1996.
46. *HyperChem 7.5, a molecular modeling product from Hypercube, Inc.* , 1115 NW 4th St., Gainesville, FL 32601
47. J. L. Duncan, D. C. McKean and P. D. Mallinson, *J. Mol. Spec.* **45,** 221 (1973).
48. W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Ab initio Molecular Orbital Theory* (Wiley, New York, 1986).
49. R. M. Silverstein, G. C. Bassler and T. C. Morill, *Spectrometric Identification of Organic Compounds,* Fifth edition; John Wiley & Sons, Inc. New York, 1991.
50. L. Horny, N. D. K. Petraco, C. Pak and H. F. Schaefer, *J. Am. Chem. Soc.* **124,** 5861 (2002).
51. R. Dembinski, et al. *J. Am. Chem. Soc.* **122,** 810 (2000).
52. http://www.denofinquiry.com/pdf/doi_HangingSlinky.pdf
53. http://scienceworld.wolfram.com/physics/SpringsTwoSpringsinSeries.html

CHAPTER 4

Modeling the Mechanical and Electronic Properties of Polyyne Molecules: A Density Functional Theory Approach

Introduction

Straight chain unsubstituted polyyne molecules are among the simplest compounds in organic chemistry [1, 2] and have been reported to be one of the constituents of interstellar clouds [3-12]. This class of compounds and their derivatives have attracted much interest in recent past [4, 13-18], due, largely to their unique electronic and optical properties [2, 4, 13-16, 18-20]. These properties make these compounds find potential applications in molecular electronics [14, 21], nonlinear optics [14] and nanotechnology [22]. By the same token, they have been reported to have potential applications as photonic devices and as efficient and precise molecular sensors [14, 17]. These compounds have also been reportedly found in numerous natural products and are suggested to have a wide range of potential applications in both biology and material science [16]. Various research groups have thus explored all these properties and potential applications over the years. These research efforts have come from the duo of experimental and theoretical fronts.

On the experimental front for example, Tykwinski's group has carried out synthesis and characterization of these compounds [2, 15, 20, 23]. They used various techniques such as ^{13}C NMR spectroscopy, differential scanning calorimetry, mass spectrometry, X-ray crystallography and UV-vis spectroscopy [15], for characterizing different properties of the compounds. Grutter and coworkers examined the electronic transitions of some of these molecules [6], while Kloster-Jensen et al. have also reported the preparation and spectral characterization of some members of this class of compounds [1]. On the synthetic side, apart from the work of Tykwinski cited earlier, others have also synthesized some of the compounds [4, 16]. Szafert and Gladysz have done a good review of the compounds already synthesized and crystallographically characterized [24].

The properties of these compounds have also been explored theoretically, including those not accessible experimentally. In this regard, Fan et al, used Restricted Hartree-Fock (RHF) *ab initio* calculations to predict the geometric and electronic structures, as well as total and binding energies of some straight chain polyynes (HC_nH; n=6-10) [25]. He used both double-zeta (DZ) and double-zeta plus polarization (DZP) contracted basis set in this study. Nalwa and coworkers conducted *ab initio* calculations to examine the effect of the nature of the π-bonding sequence on polarizabilities and second hyperpolarizabilities of buta-1, 3-diyne using STO-3G basis set [17]. In their approach, Toto et al used Møller-Plesset perturbation theory (MP2) to study the effect of electron correlation on second-order hyperpolarizability of polyynes and compared these results with that of Hartree-Fork calculations [18]. Further similar and related theoretical calculations have been carried out on these compounds by other researchers [5, 7, 9, 13, 14, 22]. Molder and coworkers have conducted a study on the molecular geometries of linear polyyne chains using PM3 method, ab initio linear combination of atomic orbitals – self-consistent field (LCAO-SCF) molecular orbital approach and density functional theory (DFT) at different basis set levels [14]. They also investigated bond length alternation in linear polyynes. Their data indicate that DFT methods give more accurate results for total energies and bond lengths than the RHF method. In a similar effort, Horny and coworkers performed calculations to predict geometries and total energies, among others, of long linear polyynes [5].

A critical look at all the above theoretical studies indicates that all of the researchers only explored the properties of unperturbed straight chain polyynes. But it has been widely reported that these molecules have a variety of applications in high technology industries, e.g. molecular electronics and nonlinear optics. Thus, if these molecules can be used as molecular wires and sensors for instance, they may have to be perturbed in a way to suit the desired purpose. Also, the fact that these molecules are conductors means that they may be amenable to interrogation with conductive atomic force microscopy (c-AFM) studies. In addition, when these molecules are bound to the surface, their structure and morphology can be characterized by AFM. Thus, these molecules may be distorted or bent under the influence of an AFM. This therefore means that apart from theoretical predictions of the properties of linear unperturbed molecules, theoretical studies of the behaviour and properties of the distorted, bent or perturbed molecules are necessary. The work reported here thus adds another dimension to the already existing knowledge on polyynes by looking at the possible behaviour of these molecules when they are bent. To our knowledge, this is the first report on theoretical studies of the mechanical and electronic properties of bent or distorted polyyne molecules.

In this work, we used computational approach to probe the possible mechanical and electronic properties of polyynes. We used various levels of theory, from semi-empirical to *ab initio* / DFT for these calculations. We monitored how the total energy changes with respect to bending, as well as the variation of total energy with stretching coordinates. The bending calculations allowed us to make reasonable predictions of how these molecules might respond to deformations that can be caused by AFM. The responses of these molecules in terms of how the relative force constants with which they might push on an AFM tip or be bent by the tip were found.

The PM3 bending calculations were carried out on C_4-C_{20}, while the stretching calculations were carried out on C_4-C_{40} members. As for the DFT calculations, we only considered C_4-C_{12} members for both bending and stretching calculations. From the resulting energy plots, we obtained force constants, which are characteristic of molecular chain lengths. It was also found that the manners in which these molecules are bent by these two methods are different, with the *ab initio*/ DFT method giving a more uniform bending across the entire chain.

From the DFT bending calculations, we also obtained the HOMO-LUMO Energy gap (Egap) of these molecules and explored its variation with bending angle. The Egap varies with chain lengths and it was discovered that the changes observed in the Egap are independent of changes in equilibrium lengths during bending. In addition, the plot of normalized Egap vs bending angle shows no difference between carbon chain lengths, as the curves essentially overlap one another.

Calculation method

All semi-empirical PM3 calculations were carried as implemented in HyperChem software [26], while all DFT calculations were performed with the GAUSSIAN 03 suite of programs [27], at the computing facility of Dr. Alex Brown of the Department of Chemistry, University of Alberta [28], and WestGrid facility [29, 30]. The machines used for the calculations were x86-Linux computer [28] and Tru64 Alpha systems [29].

The PM3 calculations were automated by using the Dynamic Data Exchange (DDE) implemented in HyperChem™ [26]. By this means, we were able to issue commands from Excel™ application to drive the computations in HypeChem™. We wrote a series of macros, which enabled all the computations to be automated such that the computations in HyperChem™ were driven from an Excel™ application and the results were immediately recorded in real time in a spreadsheet. We then encode input job files in scripts that were subsequently submitted to the workstations on the to perform the DFT computations.

Based on our previous related calculations [31], and some published works in the literature [32-38], we used BLYP functionals [39, 40-41] with the large basis set 6-31G**. For the DFT calculations, we started with rough geometry optimizations using semi-empirical PM3 method as implemented in HyperChem™ software [26], the resulting atomic coordinates and geometries were then used as starting points in the more accurate DFT calculations. Geometries of the molecules were fully optimized with tight convergence criteria and with ultrafine numerical integration pruned grids (99 radial and 590 angular points).

The model used for the bending calculations (see Figure 4.1) is such that the terminal C1-H1 bond was constrained so that the two atoms do not move at all during the optimization, while the x-y-coordinates of the terminal H2-atom were varied to bend the molecule through a desired angle in 10^{o} increments from $0\text{-}90^{o}$. At each bending angle, the minimum total energy was obtained via geometry optimization. Once the molecule is bent, the coordinates of the terminal H2-atom were also allowed to vary along the straight line defined by the bending angle to attain the minimum energy possible for that angle. We obtained the minimum total energy for each geometry optimized molecule at each bending angle and then plotted the energy as a function of bending angle. For the energy curves, we used offset total energy, which was obtained in each case by subtracting the total energy of straight unbent molecule (i.e. at 0^o) from that at a given bending angle. The bending angle was converted to "equivalent length" $\boldsymbol{l}$ in order to easily obtain force curves and force constants from the bending calculations.

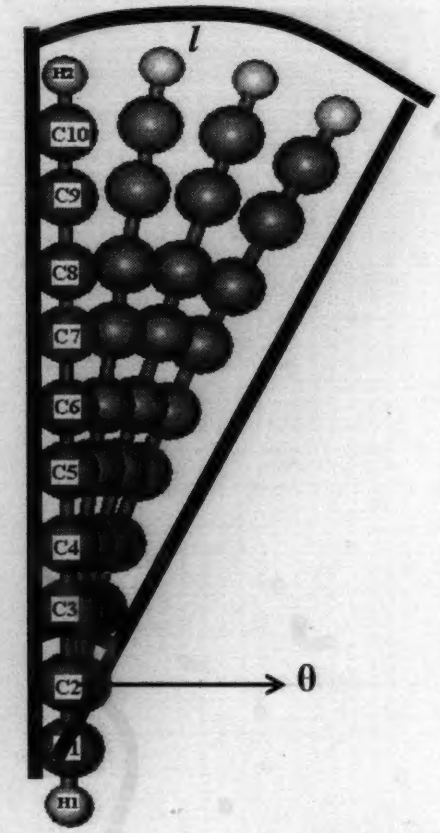


Figure 4.1: Model of a polyyne molecule bent through 30°. The terminal C1-H1 bond was constrained while the x-y-coordinates of the terminal H2-atom were varied to bend the molecule in 10° increments from 0-90°. The applied constraint effective to ensure that the fixed atoms do not essentially move during all geometry optimization processes was 10,000kcal/molÅ^2 (7000N/m) and this does not significantly affect the calculated total energy. The 10° increments gave essentially same results with those simulations carried out at 1° increments from 0-90°. In each case, the minimum total energy was obtained via geometry optimization.

Equivalent length ***l*** described when the molecule is bent through an angle **θ** is obtained in segments for each 10° increment in bending angle up to 90°. For a particular bending angle, the arc length is given by:

$$Arc = (\theta/360) * 2\pi R$$

where θ is the bending angle in degrees

R is taken as the average of the two "bracketing radii" since the two Rs are unequal as in a perfect circle and it is in Å.

The cumulative arc length at a bending angle is termed the equivalent length ***l*** for that angle. For example, the equivalent length at 20^0 bending angle is found by adding the arc length at 10^0 to that at $20^0$and so on. This equivalent length is used in the energy and force curves as it permits direct calculation of force constants. Percentage equivalent length used in some instances is given by:

$$\%E_l = (E_l / O_l) * 100\%$$

where E_l and O_l are equivalent length and original length respectively

For stretching (Figure 4.2), both H-end-atoms were constrained, while the x-coordinate of the H2 end atom was changed along an axis. Each molecule was stretched in 0.1Å increments up to 1Å, and the minimum energy was also obtained by geometry optimization. The applied constraints effective to ensure that the fixed atoms do not move at all during all geometry optimization processes were 10,000kcal/molÅ^2 (7000N/m).

The % stretched length **%(ΔL / L$_0$)** used in some plots is given by:

$$\%S_l(\%\nabla L / L_o) = (\nabla L / L_o) * 100\%$$

where S_l is the stretched length

∇L = L$_l$- Lo, Lo is the original unstretched length and **L$_l$** is the stretched length.

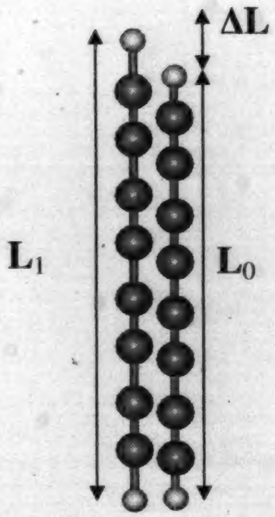


Figure 4.2: A cartoon of the stretching simulation. The unstretched length $\mathbf{L_0}$, the stretched length $\mathbf{L_1}$ and the length increment $\mathbf{\Delta L}$ are shown. For the stretching simulations, both H-end-atoms were constrained, while the coordinate of the H2 end atom was changed along x-axis. Each molecule was stretched in 0.1Å increments up to 1Å, and the minimum energy was obtained by geometry optimization. The applied constraint effective to ensure that the fixed atoms do not essentially move during all geometry optimization processes was 10,000kcal/molÅ^2 (7000N/m) and this number does not significantly affect the calculated total energy.

Energy calculations on surface-bound 1, 3, 5, 7, 9- decapentayne were also carried out to see how different the behaviour of the molecule would be compared to a free unbound molecule with both PM3 and DFT methods. In this model, one terminal carbon atom was bonded to one silicon atom of a slab consisting of four silicon atoms fixed at the ideal Si (111) lattice positions. Such a small number of Si atoms was used to ensure reasonable computation time. The silicon dangling bonds were capped with hydrogen atoms. The bending calculations were performed in an analogous way to the free molecules by fixing the terminal C1-Si bond. However, calculations performed with the C1-Si bond unrestrained gave the same force constants.

Force curves were obtained as first derivatives of energy curves, while force constants were obtained as second derivatives. Since total energy is in units of kcal/mol and equivalent length is in angstroms, then the first derivative of energy vs equivalent length curves will be kcal/molÅ and the second derivative will be in kcal/molÅ^2. Thus, all force constants k in kcal/molÅ^2 were converted to N/m using the conversion factor 1 kcal/molÅ^2 = 0.7N/m [42].

To gain insight into the bending impact on their electronic properties, we used the DFT approach for electronic structure calculations (i.e. HOMO-LUMO Energy gap (Egap) calculations). In this case, we normalized the Egap at a particular bending angle to that at 0°, and we used the average atomic angle as a means of normalizing the bending angles for all molecules. By average atomic angle, we mean the average bending angle observed for all the carbon atoms in a molecule at a particular bending angle.

Results and Discussion

PM3 Bending Calculations

Energetics

For energy calculations, we explored how the total energy of these molecules changes with respect to bending. We found that based on chain lengths, the slopes of the energy curves are different indicating different force constants for the molecules. The energy curves obtained from the results are presented below.

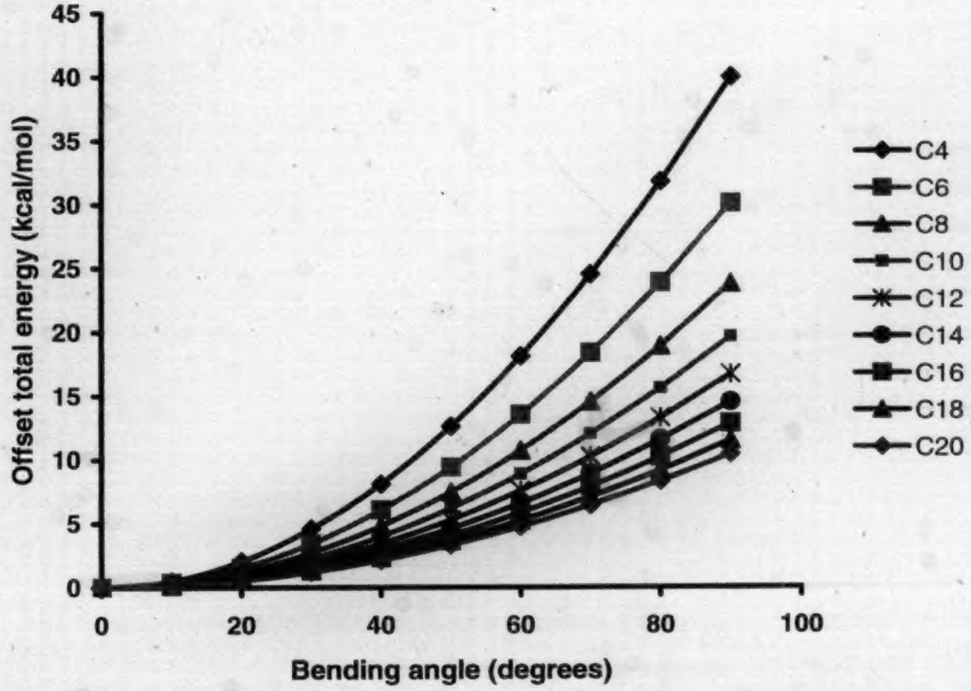


Figure 4.3: Plots of offset total energy vs bending angles for C_4-C_{20}. The offset total energy was obtained in each case by subtracting the total energy of straight unbent molecule (i.e. at 0^o) from that at a given bending angle.

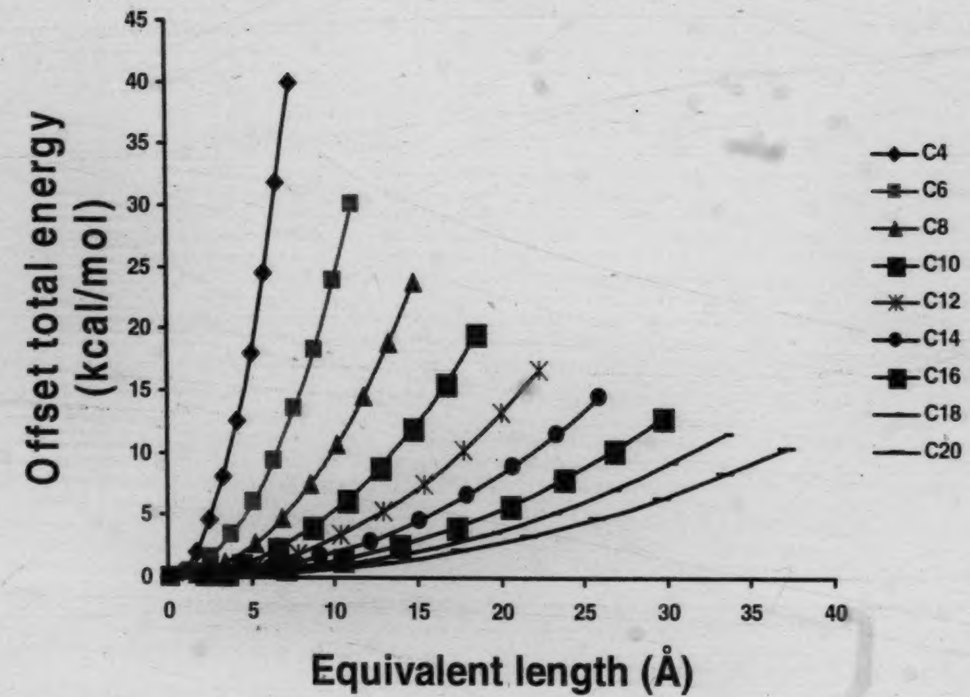


Figure 4.4: Plots of offset total energy vs equivalent length for C_4-C_{20}. These plots have same values on the y-axis as Figure 4.3 above while the bending angles (in degrees) in Figure 3.3 are converted to equivalent length in Å. The equivalent lengths are as explained previously.

For surface-bound 1, 3, 5, 7, 9- decapentayne, our bending simulations revealed a difference in the total energy of this molecule, compared to the free unbound molecule. This is not unexpected, given the fact that the silicon atoms and the saturating hydrogen atoms would contribute to the total energy. The force constant obtained for the free molecule and the surface-bound one are respectively 0.085 N/m and 0.05 N/m.

We also examined how the bending occurs along the entire chain of each molecule. Figure 3.6 shows the plots of C-C-C bond angles for C_4-C_{20} bent at 20^o. From the curves, it can be seen that for each molecule, the bond angles generally increase from C_1-C_2-C_3 to C_2-C_3-C_4 and so on, meaning that the bending is largest at the C_1-C_2-C_3 bonds and decreases further along the chain. In addition, the overall extent of bending decreases with increasing chain lengths. This is evident considering the slope of each curve. This observation can be tied to the chain length differences of the molecules. Based on length, longer chains are less bent compared to shorter chains at the same bending angle.

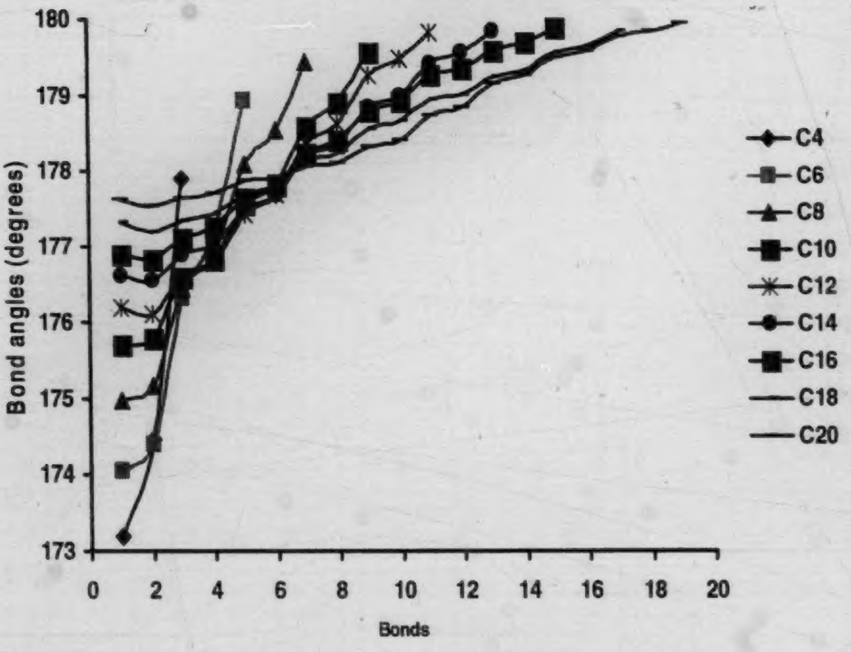


Figure 4.5: Plots of C-C-C bond angles for C_4-C_{20} bent at 20°. It can be seen from the plots that for each molecule, the bond angles generally increase from C_1-C_2-C_3 to C_2-C_3-C_4 and so on, indicating a decreasing level of bending as the chain is traversed. The overall extent of bending decreases with increasing chain lengths as can be seen from the slope of each curve.

Force Curves

From the energy curves, we obtained force curves and force constants for the molecules as presented in the following figures. The force constant curve shown in Figure 4.7 was fitted with the function derived in chapter 4 for bending of polyynes, i.e.

$$K_{eff} = (2/k_1 + ((n-4)/2)/k_2)^{-1}$$

With this fitting function, the k_1 and k_2 force constants became fitting parameters as the constant values changed to 2.10 and 0.39, respectively. However, using this function to fit the curve attaches some physical meaning to the bending of these molecules. The discrepancy between these fitting parameters and the force constant values is perhaps due to the fact that our approximation of the polyyne molecules as units of springs joined together in series is invalid for bending of these molecules. The root of the approximation lies in stretching of springs joined together in series and not in bending them.

We also used another fitting function which is based on the approximation that the molecules can be regarded as flexible rods. The function, which is based on the flexural rigidity of rods [43] is given by:

$$k = 3EI / L^3$$

where EI is the term called flexural rigidity of the rod
E is the Young's Modulus of the rod
I is the moment of inertia
L is a measure of the length of the rod.

If the polyyne molecules are approximated as rods, then their bending can be related to the above expression based on flexural rigidity. The values of E, I and L from our fit are respectively 3.35 x 10^3Nm^{-2}, 0.01m^4 and 0.61Å. The values of E and I are however different from their original constant values which are respectively 4.7 x

10^{13}Nm^{-2} and 0.9 x 10^{-42}m^{4} [44]. The value of L is variable as it is a measure of the length of the molecule which is different from one molecule to another. The difference between our fitting parameters and the constant values of E and I may be ascribed to the fact that our approximation of the molecules as flexible rods is less adequate, or it may be because the original constant values stated above are also based on computations and mechanical engineering approximations [44], which may not be adequate. However, as can be seen in Figure 4.7, the flexural rigidity fitting function gives a better fit than the K_{eff} function.

It should be noted that both functions fit the force constant curves reasonably well and that the fitting parameters have physical meaning. These two functions are used to fit all the subsequent force constant curves.

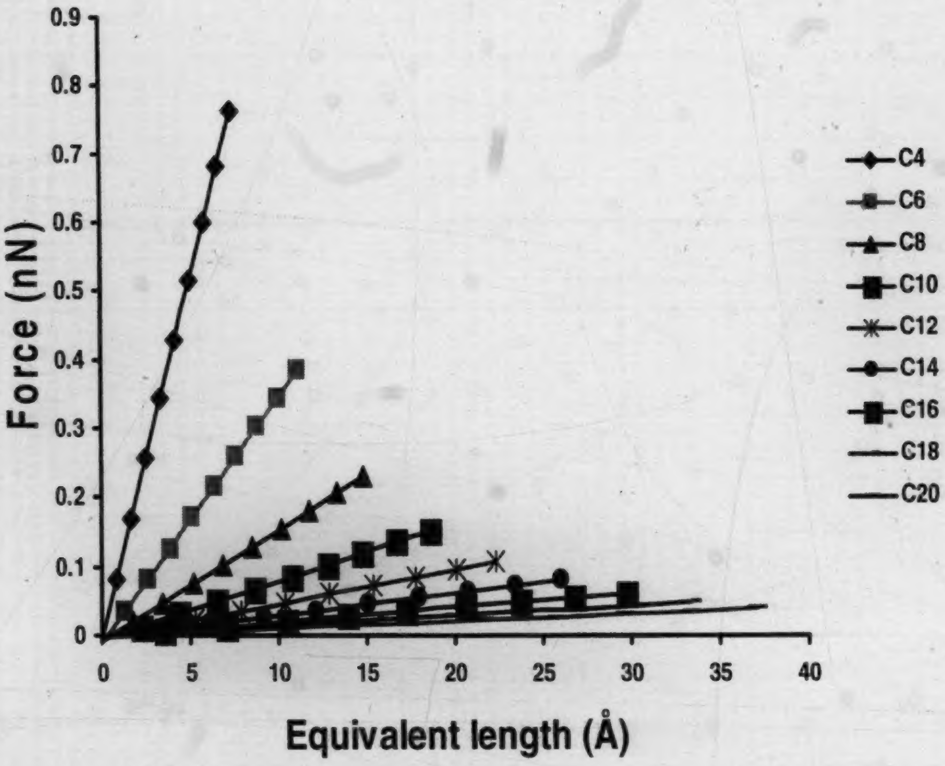


Figure 4.6: Plots of force vs equivalent length for C_4-C_{20}. The force was obtained as the first derivative of the energy curves.

# of carbon atoms	k (N/m)
4	1.04
6	0.356
8	0.16
10	0.085
12	0.05
14	0.032
16	0.022
18	0.015
20	0.011

Table 4.1: The force constants obtained from the bending of selected polyynes. Force constants of some of these molecules may fall in the range of some AFM cantilevers' force constants depending on their chain lengths.

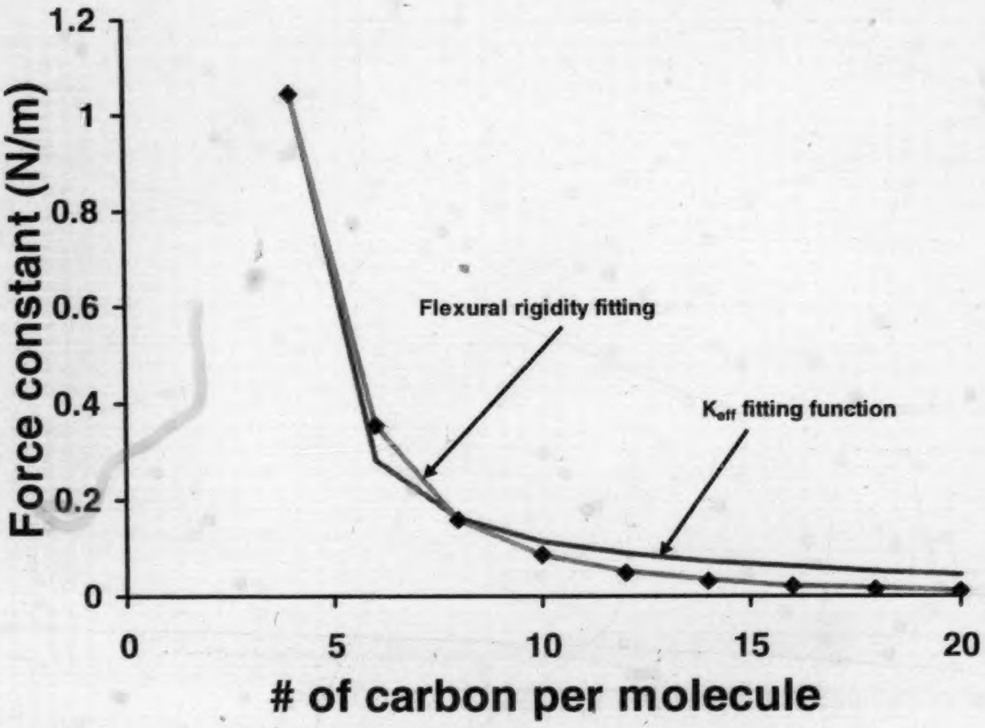


Figure 4.7: Plot of force constants vs number of carbon atoms per molecule of polyyne. From this curve, it can be seen that the force constant k decreases as the chain length increases. The fitting functions are also shown. The K_{eff} function does not fit the data as much as the flexural rigidity function does, showing the latter function may be more suitable for bending of polyyne. The K_{eff} function is $K_{eff} = (2/k_1 + ((n-4)/2)/k_2)^{-1}$, while the flexural rigidity function is $k = 3EI/L^3$.

PM3 Stretching Calculations

Energetics

In a fashion similar to the bending simulation, we explored the variation of the total energy of these molecules with respect to stretching coordinates. To this end, we performed calculations on C_4-C_{40} members, as explained above. The resulting energy curves are presented below. Stretching force constants were obtained from the energy or force curves in an analogous way to the bending.

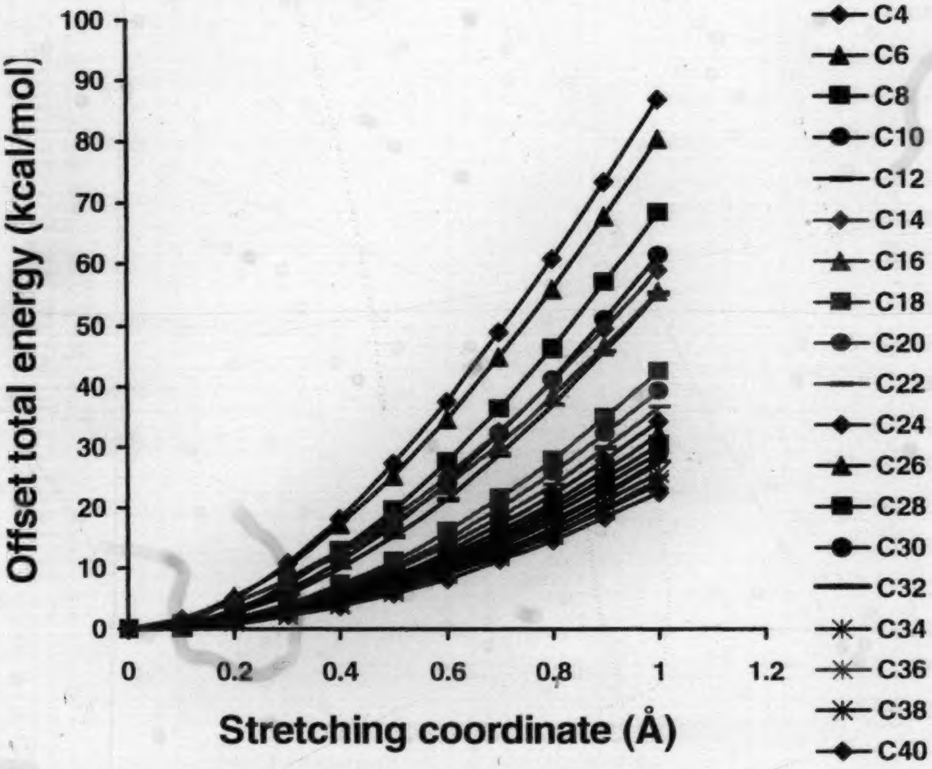


Figure 4.8: Plots of offset total energy vs stretching coordinate for C_4-C_{40}. These curves follow the same trend as for bending energy curves. The slope of curves decrease as chain lengths increase, signifying decreasing force constants with chain length increase.

Force Curves

Stretching force constants were obtained as second derivatives from the energy curves. These stretching force constants are about two orders of magnitude higher than for bending force constants, indicating that a higher energy is needed for stretching than for bending. This is similar to the relationship observed between bending and stretching force constants that we obtained from the experimentally observed wavenumbers of DP-C12 in chapter 3.

The fact that the stretching force constants are higher than the bending force constants is in agreement with a similar phenomenon observed with vibrational modes of molecules, where the force constants observed for bond stretching modes are higher than those of bending modes, indicating the higher energy required for stretching than for bending.

The force constant curve was fitted with the function

$$K_{eff} = (2/k_1 + (n/2)/k_2 + ((n/2)-1)/k_3)^{-1}$$

which was derived in chapter 3 for stretching of polyynes. The force constants k_1 k_2 and k_3 respectively have values of 226.60, 1776.82 and 1776.90, which are slightly different from the constant values originally obtained in chapter 3. This difference might imply that our approximation is less adequate, but since these constant values are closer to the original values (especially for k_2), the approximation seems more adequate for stretching than for bending. Also, the fact that k_2 and k_3 have almost exactly same values indicate that the molecules are sufficiently conjugated across the entire chain lengths, i.e. no separate single and triple bonds. The flexural rigidity function was also used to fit the curve with E, I and L values of 4.93 x 10^3Nm^{-2}, 2957.42 m^4 74.03Å, respectively. The force constants obtained from our computations are shown in the following table and curve.

Carbon #	k (N/m)
4	90.6
6	84.1
8	82.2
10	76.3
12	70.4
14	61.8
16	57.3
18	56
20	52.2
22	48.8
24	45.9
26	43.2
28	40.8
30	38.6
32	36.7
34	35
36	33.4
38	31.8
40	30.6

Table 4.2: The stretching force constants obtained for C_4-C_{40} polyynes. Values are about two order of magnitude times those obtained for bending simulations and are seen to decrease with chain lengths.

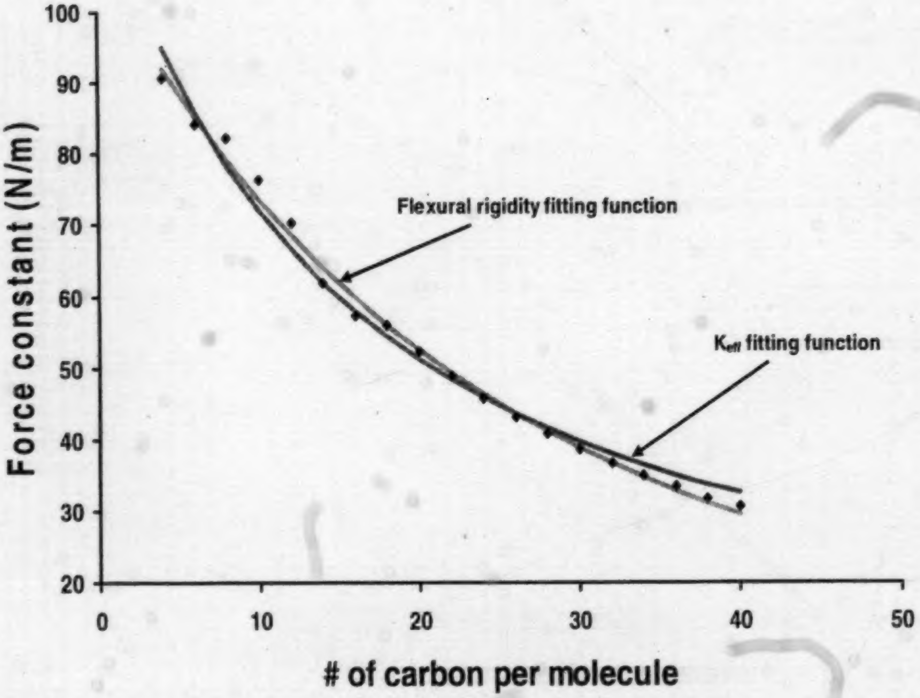


Figure 4.9: Plot of force constant vs number of carbon atoms per molecule for C_4-C_{40}. Force constants were obtained as second derivatives of the corresponding energy curves and are about two orders of magnitude times those obtained for bending simulations. The data was fitted with K_{eff} fitting function $K_{eff} = (2/k_1 + ((n-4)/2)/k_2)^{-1}$, and the flexural rigidity fitting function $k = 3EI/L^3$.

To reveal the intrinsic differences between these molecules, we obtained a plot of stretching force vs % stretched length. By normalizing the stretching coordinates of each molecule to its unstretched length this way, the plot enabled us to see the intrinsic differences between the members of polyyne series considered, apart from chain length differences. We can see the variation of the stretching force as a function of strain for each molecule.

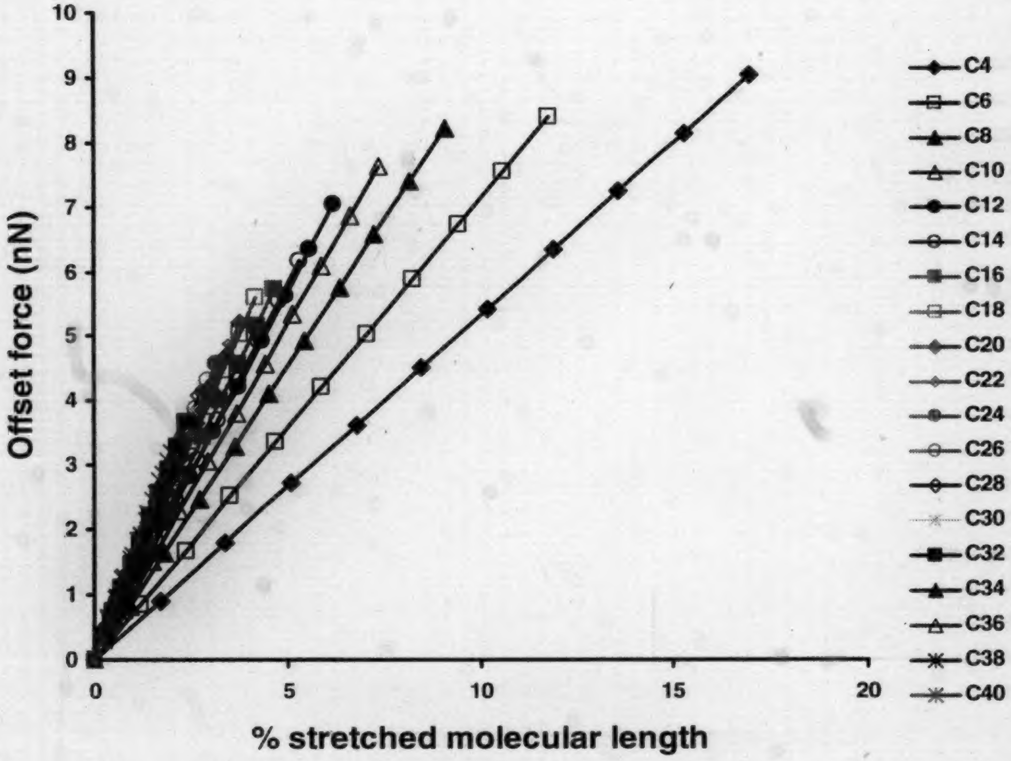


Figure 4.10: Plots of offset force vs % stretched molecular length $(\%\nabla L / L_o)$ for C_4-C_{40}. These plots reveal the intrinsic differences between these molecules other than chain length differences.

To further reveal these differences, we obtained the slope of the curve in Figure 4.10 and plotted it against the number of carbon atoms per molecule. The resulting curve (which is analogous to Young modulus) shows an interesting trend in the intrinsic mechanical property of these molecules. This property, as shown in the following curve, is almost opposite to the force constants and may attain a limiting value when the molecule is rather infinitely long.

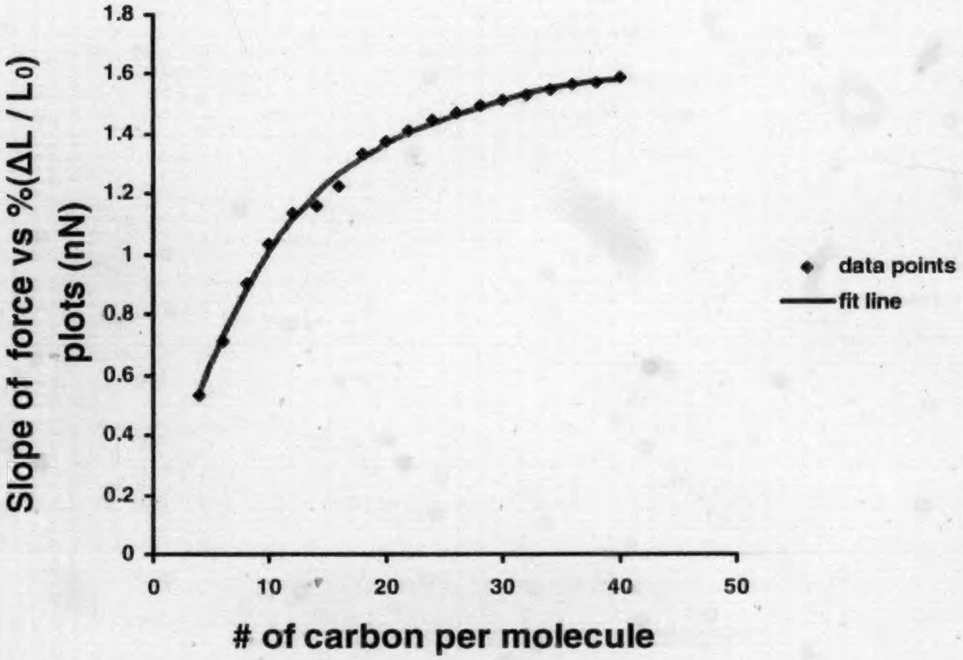


Figure 4.11: Plot of the slope of the plots shown in Figure 4.10 vs number of carbon atoms per molecule for C_4-C_{40}. This plot shows an interesting trend in the intrinsic mechanical property of these molecules, which is different from force constants shown in Figure 4.9 above. As seen, the slope of this curve is almost opposite that of the force constant curve in Figure 4.9.

DFT Bending Calculations

Energetics

For the DFT bending simulations, we monitored how the total energy changes as a function of bending angle for C_4-C_{12}. The energy curves are shown in Figure 4.12. The total energies obtained by our method for the straight unbent molecules are comparable to reported values to within 0.04% [5, 14]. It must be mentioned that the bending angles of 0-90° may be rather extreme, for in reality the molecule might not be bent to beyond 30° under the influence of an AFM tip. Thus, the calculations carried out for the surface-bound molecule was only performed from 0-40°.

Similar to the case of PM3 calculations, our bending simulations on surface-bound 1, 3, 5, 7, 9- decapentayne revealed a difference in the total energy of this molecule, compared to the free unbound molecule, giving a force constant of 0.06 N/m for the free molecule and 0.012 N/m for the surface-bound. This difference is due to the contributions of the silicon and hydrogen atoms constituting the surface to the total energy of the system.

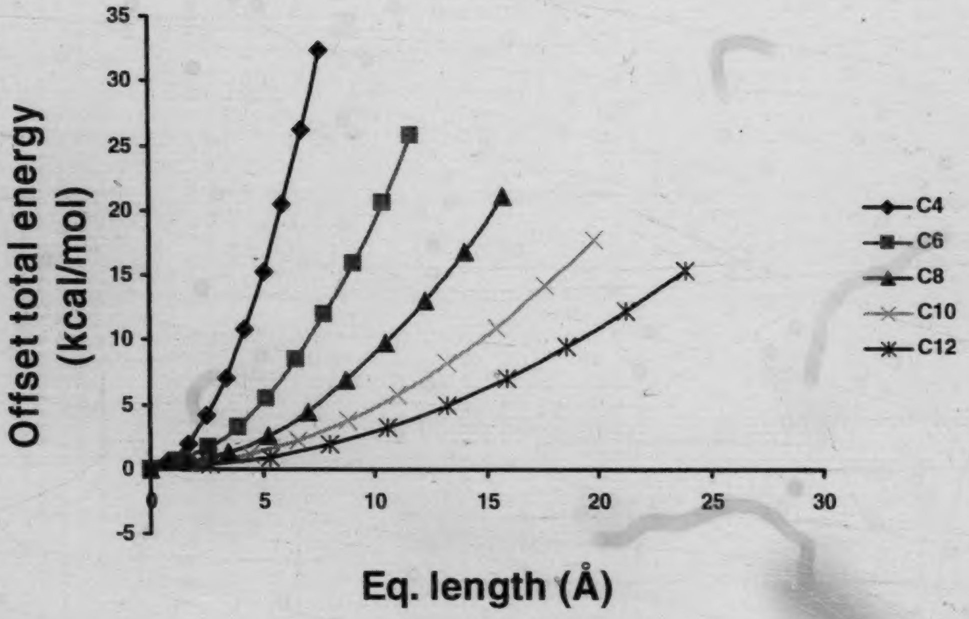


Figure 4.12: Plots of offset total energy vs equivalent length for C_4-C_{12} polyynes. These plots were obtained from our DFT calculations and so are expected to be more accurate than the corresponding ones from PM3 calculations.

Comparing the manner in which the molecules are bent in PM3 and DFT, we also found that for PM3 optimization, most of the bending is concentrated on the C_1-C_2-C_3 bonds while in DFT the bending is uniformly spread across the entire chain length (Figure 4.13). The fact that the bending is more pronounced at the C_1-C_2-C_3 for PM3 bending has been stated in Figure 4.5 when comparing the extent of bending along the molecular chain for PM3 calculations. In both cases however, the bending decrease further along the chain, but gradually so in the case of DFT simulations. This becomes more evident considering the C-C-C bond angles across the chains as represented in Figure 4.14.

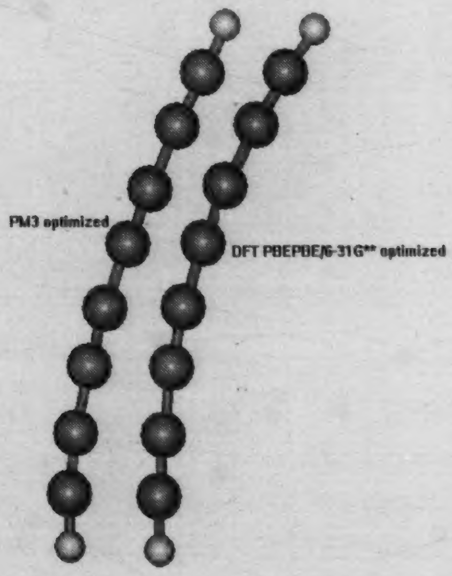


Figure 4.13: A cartoon of 1, 3, 5, 7-octatetrayne molecules bent at 20° and geometry optimized by PM3 and DFT. It can be seen from this figure that most of the bending is concentrated on the C_1-C_2-C_3 bonds for PM3 method, while DFT tends to cause bending of the molecule in a uniform manner across the entire chain.

Similarly, for our DFT calculations, we examined how the bending occurs along the entire chain of each molecule. Figure 4.14 shows the plots of C-C-C bond angles for C_4-C_{12} bent at 20^o. Though the curves show a rather complicated pattern, it can be seen from these curves that for each molecule, the degree of bending is highest at the C_1-C_2-C_3 bonds and decreases gradually further along the chain. In addition, the overall extent of bending decreases with increasing chain lengths. This is evident considering the slope of each curve and the C-C-C bond angles. This is due to the chain length differences of the molecules. At a given bending angle, longer chains appear to be less bent compared to shorter chains based on chain length differences.

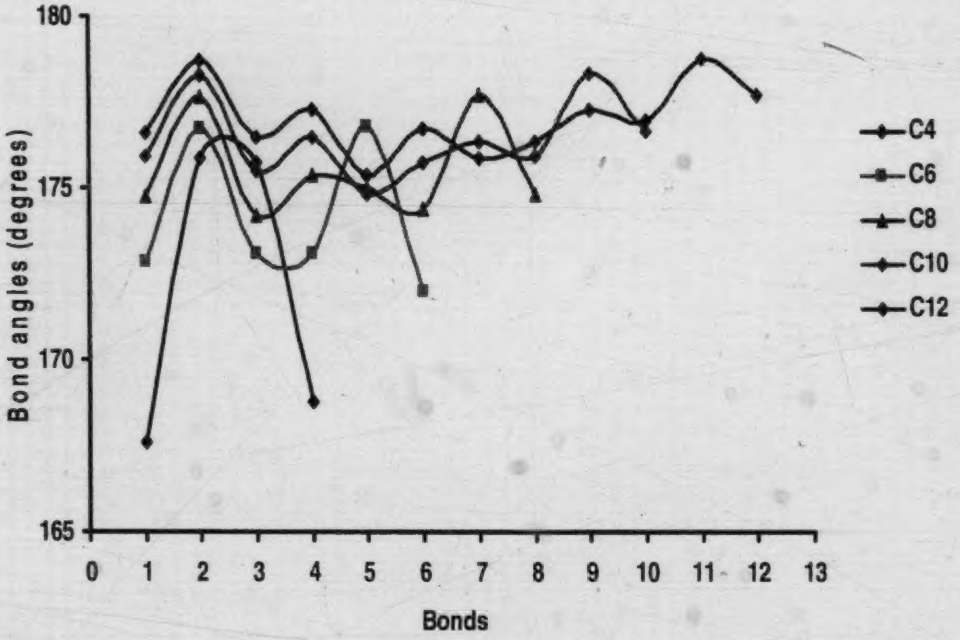


Figure 4.14: This figure shows the plots of C-C-C bond angles for C_4-C_{12} bent at 20°. The bond angles are seen to be within the range shown for PM3 calculations (Figure 4.5 above). Thus, the manner in which the molecules are bent in this case is such that the greatest bending occurs at C_1-C_2-C_3 bonds as in PM3 simulations, but gradually decreases further along the chain, unlike in PM3 simulations where the decrease is rather drastic. As can be seen from the plots, for each molecule, the bond angles oscillate from C_1-C_2-C_3 through to Cn_{-2}-Cn_{-1}-C_n as the chain is traversed from C_1 to C_n. The overall extent of bending decreases with increasing chain lengths as can be seen from the slope of each curve and the bond angles.

Force Curves

We obtained force curves and force constants from the energy curves. It is seen that the force constants decrease as the chain lengths increase. The force constants obtained, as seen in Figure 4.15, may fall in the range of some cantilever's force constants depending on chain lengths, thus allowing reasonable predictions to be made about how these molecules might respond to deformations that can be caused by AFM cantilevers. Also, the curve was fitted with same functions as were used for the PM3 curves. The fitting parameters for the Keff function k_1 and k_2 are 1.43 and 0.30, respectively. These values are also significantly different from the original force constants obtained in chapter 3, again this may be because the approximation that these molecules behave as springs joined together in series is inappropriate for these systems. For the flexural rigidity function the E, I and L values are respectively, $3.35 \times 10^3 \text{Nm}^{-2}$, 0.01m^4 and 0.69Å. Again, the discrepancy observed in these parameters may be due to the fact that our approximation that these molecules behave as rods is less adequate, or may be because the constant values are also results of theoretical calculations and mechanical engineering approximations, which in themselves may be inadequate.

# of carbon atoms	**k (N/m)**
4	**0.712**
6	**0.248**
8	**0.113**
10	**0.060**
12	**0.026**

Table 4.3: This table shows the bending force constants obtained for the selected members of the polyyne series from our DFT calculations. Shorter chains are stiffer than longer ones based on the force constant values.

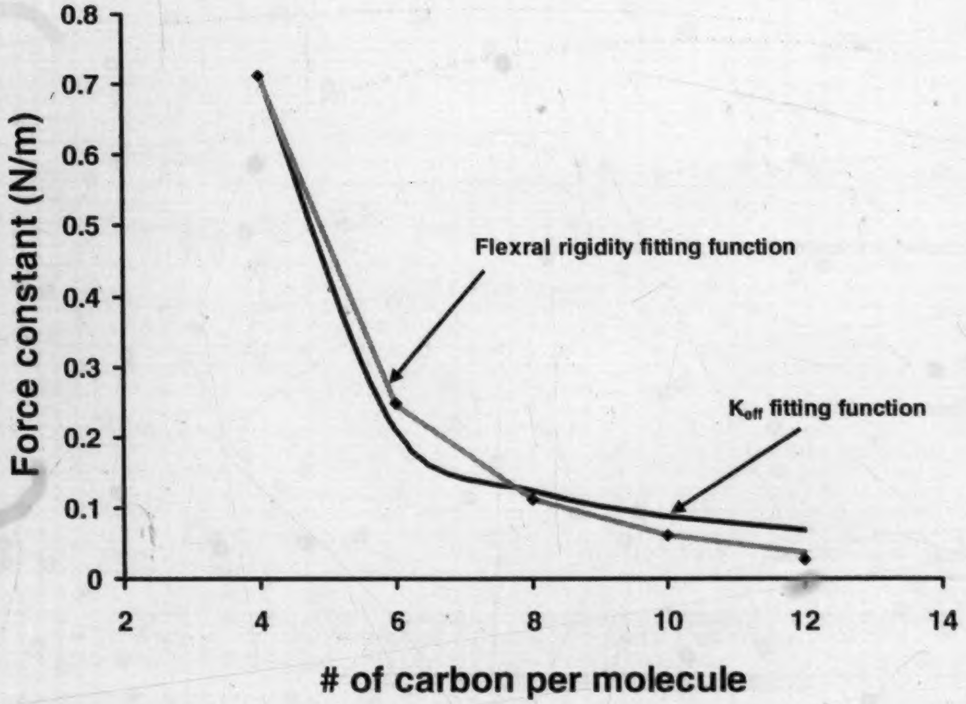


Figure 4.15: Plot of force constant vs number of carbon atoms per molecule for C_4-C_{12} polyynes. Since DFT calculations are more accurate than PM3, this curve is more reliable for estimating force constants and how these molecules might respond to deformations that can be caused by AFM cantilevers. The K_{eff} fitting function $K_{eff} = (2/k_1 + ((n-4)/2)/k_2)^{-1}$, and the flexural rigidity fitting function $k = 3EI/L^3$ are shown.

DFT Stretching Calculations

Energetics

As was done for the PM3 calculations, we obtained energy curves for the DFT stretching simulations. Force constants for the different molecules were determined from the energy curves. Due to the fact that DFT calculations are more accurate than PM3, the curves presented below should be more reliable than for the PM3 calculations.

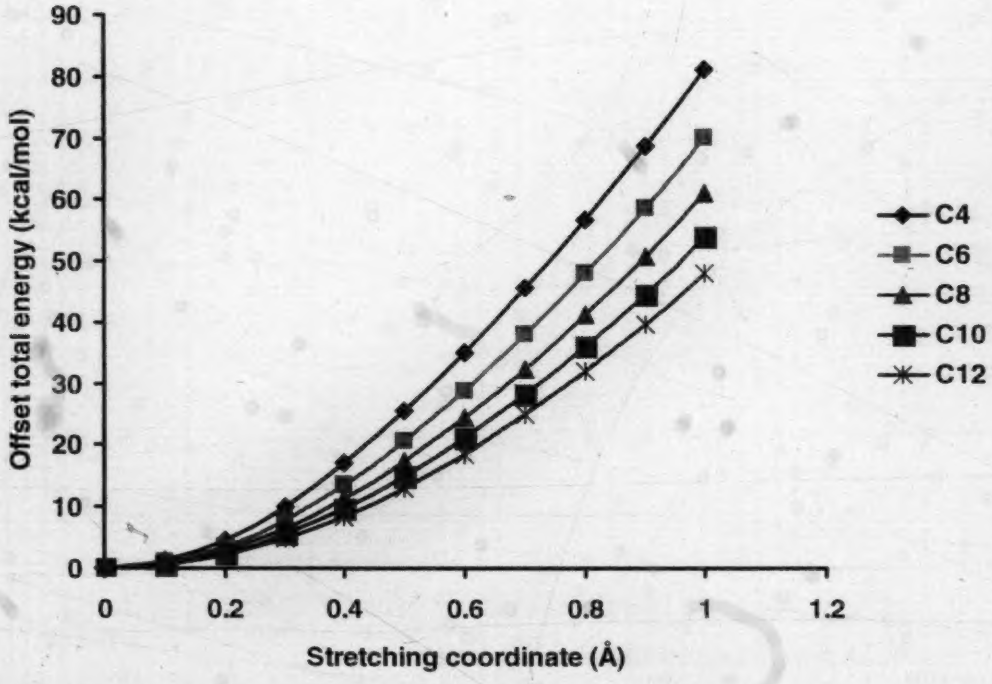


Figure 4.16: Plots of offset total energy vs stretching coordinate for C_4-C_{12}. The slopes of the curves decrease as chain lengths increase from C_4 to C_{12}, signifying that the resulting force constants are chain length dependent.

Force Curves

Force curves and force constants were obtained from the energy curves. The stretching force constants obtained from the energy curves are about two orders of magnitude times those obtained for the bending simulations. We observed similar relationship between bending and stretching force constants from the experimentally observed wavenumbers of DP-C12 in chapter 3. The force constant curve was fitted with the stretching K_{eff} function as before, and the force constant values k_1, k_2 and k_3 are 201.20, 1757.07 and 1959.52, respectively. These values are closer to the original effective stretching force constants estimated in chapter 3, when compared to the values we obtained for the DFT bending calculations. The curve was also fitted with the flexural rigidity function with E, I and L values of 4.50 x $10^3 Nm^{-2}$, 2252.90 m^4 66.70Å, respectively.

# of carbon atoms	k (N/m)
4	84.8
6	80.6
8	74
10	67.4
12	61.4

Table 4.4: This table shows the stretching force constants obtained for the selected members of the polyyne series from our DFT calculations. The force constants are chain length dependent.

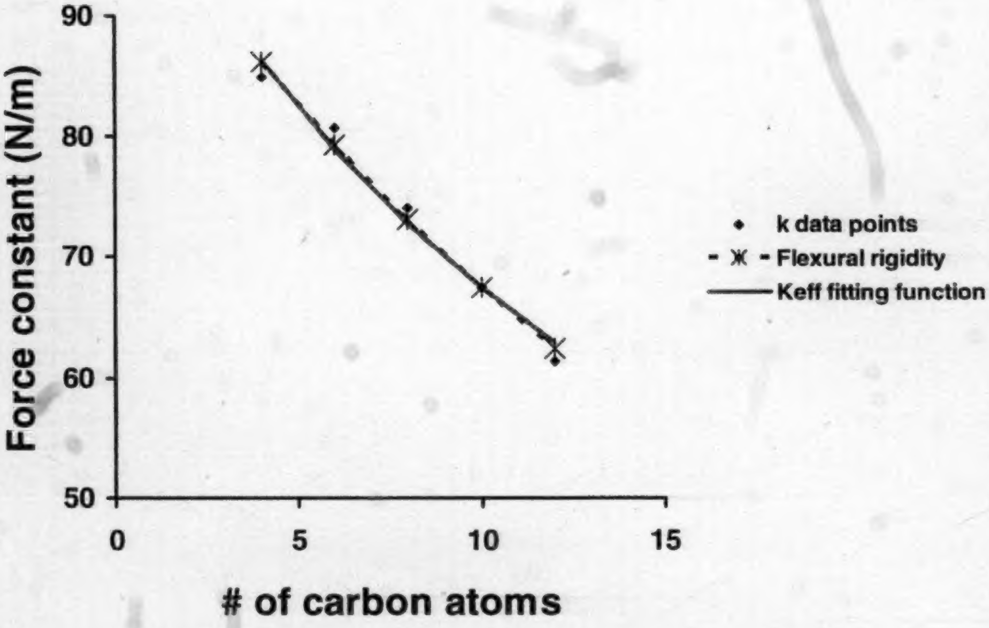


Figure 4.17: Plot of stretching force constant vs number of carbon atoms per molecule for C_4-C_{12}. Similar to the case of bending, the force constants are seen to be chain length dependent and are about two orders of magnitude times those obtained for the bending simulations (Figure 4.15). The K_{eff} fitting function $K_{eff} = (2/k_1 + ((n-4)/2)/k_2)^{-1}$, and the flexural rigidity fitting function $k = 3EI/L^3$ were used to fit the data as shown. The two functions essentially overlap each other.

We also obtained a plot of stretching force vs % stretched length. This plot enabled us to see the intrinsic differences between the members of polyyne series considered, apart from chain length differences when the stretched lengths are normalized this way.

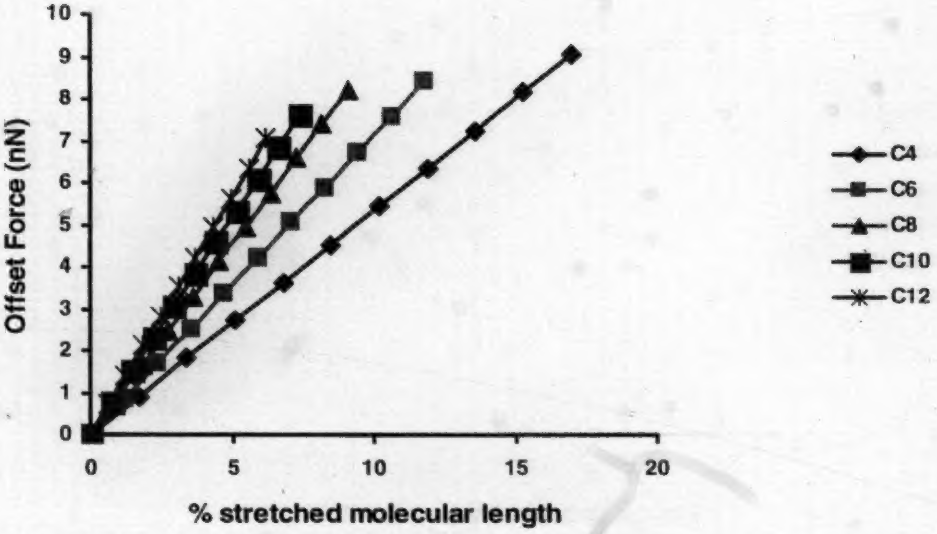


Figure 4.18: Plots of offset force vs % stretched molecular length $(\%\nabla L / L_o)$ for C_4-C_{12}. These plots reveal the intrinsic differences between these molecules other than chain length differences after normalizing the stretched lengths.

HOMO-LUMO Energy gap (Egap)

We used DFT approach to gain insight into changes in electronic properties of these molecules as a function of bending, by obtaining HOMO-LUMO Egap for these molecules at the various bending angles. We observed a decreasing trend in the Egap as a function of bending angle for each molecule. At each bending angle, we also observed that the Egap decreases with increase in chain lengths. Taking the energy levels separately, the HOMO energy levels increase with chain lengths while the LUMO energy levels decrease with chain lengths. The overall change (i.e. Egap) in each case shows a decreasing trend in Egap with chain lengths and bending angle (Table 4.5). Figure 4.19 shows the plots of normalized Egap as a function of bending angles for the five molecules considered.

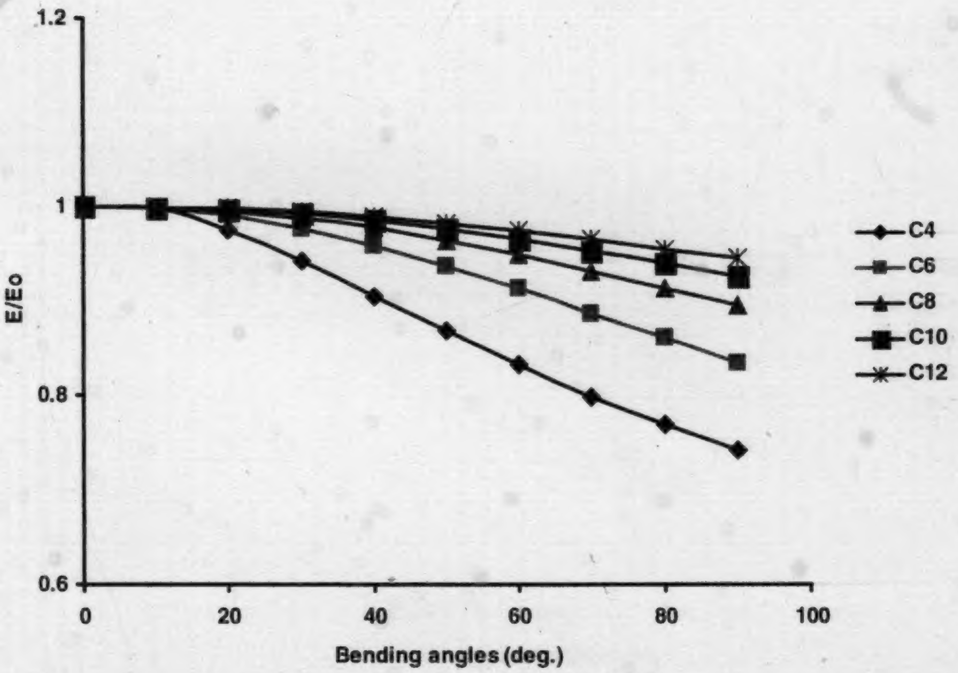


Figure 4.19: The HOMO-LUMO Egap normalized to the Egap at 0^o bending angle. The decreasing trend of the Egap as a function of bending angle is apparent. This decrease appears to manifest after bending through 10^o bending angle.

To see whether the changes in Egap with differences in chain lengths are indeed dependent on bending angles for these molecules, we obtained a plot of normalized Egap as a function of average atomic angle. By average atomic angle, we mean the average bending angle observed for all the carbon atoms in a molecule at a particular bending angle. As can be seen in Figure 4.20, the curves for all the five molecules essentially overlap one another. Thus, apart from differences in chain lengths, the extent of bending of each molecule at a particular bending angle is essentially the same for all the molecules and the decreasing trend observed with the Egaps is independent of chain lengths. We also found that the trend observed with the Egap is not dependent on the varying lengths of the molecules during bending, as we obtained no correlation between the Egap and varying lengths of these molecules.

Comparing the free unbound polyyne molecule with that bound to a silicon surface, the HOMO-LUMO Egap only differs by a constant value of about 0.005 hartrees for each bending angle (lower for the surface-bound molecule). The difference in Egap between surface-bound 1, 3, 5, 7, 9- decapentayne and unbound 1, 3, 5, 7, 9- decapentayne is indicative of the contribution of the four silicon atoms and the saturating hydrogen atoms to the total energy, which also reflects in the Egap observed, thereby influencing its electronic properties.

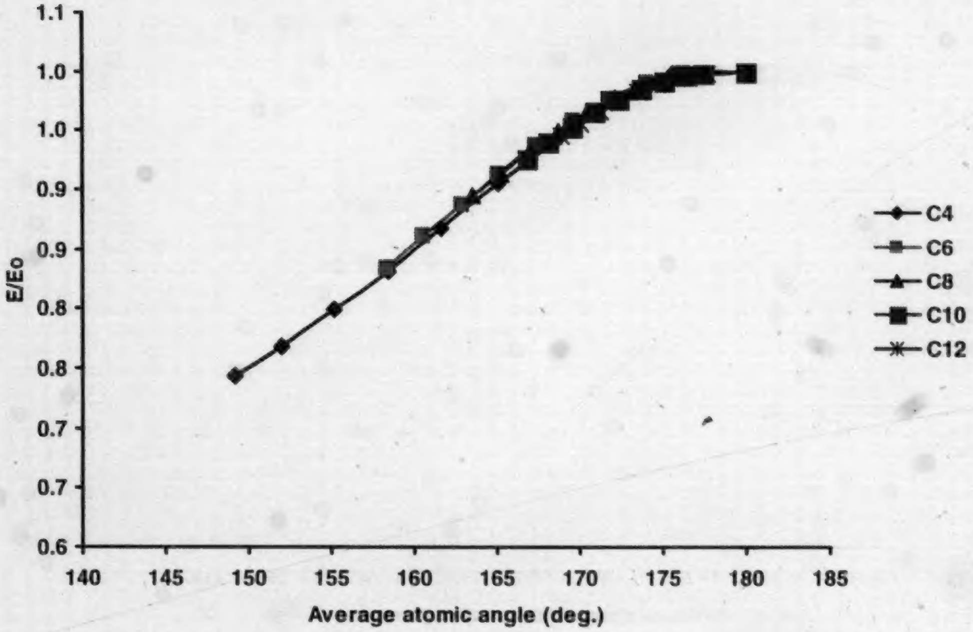


Figure 4.20: Normalized Egap plotted as a function of average atomic angle. Average atomic angle, is the average bending angle observed for all the carbon atoms in a molecule at a particular bending angle. The curves for all the molecules essentially overlap each other. From these curves, it can be seen that apart from differences in chain lengths, the extent of bending for each molecule at a particular bending angle is essentially the same for all the molecules. So, the decreasing trend observed with the Egaps is independent of chain lengths.

	Egap (hartrees)				
Bending angle	**C4**	**C6**	**C8**	**C10**	**C12**
0	**0.17066**	**0.1289**	**0.10432**	**0.08806**	**0.0765**
10	**0.17046**	**0.12862**	**0.10416**	**0.08794**	**0.0764**
20	**0.16623**	**0.12766**	**0.10383**	**0.08782**	**0.07634**
30	**0.16071**	**0.12589**	**0.10309**	**0.08739**	**0.07608**
40	**0.15443**	**0.12357**	**0.10198**	**0.08677**	**0.0757**
50	**0.14815**	**0.1208**	**0.1006**	**0.08595**	**0.07518**
60	**0.14197**	**0.11773**	**0.09903**	**0.08504**	**0.07456**
70	**0.13633**	**0.11427**	**0.09724**	**0.08395**	**0.07386**
80	**0.13117**	**0.11093**	**0.09529**	**0.08279**	**0.07312**
90	**0.12683**	**0.10756**	**0.09333**	**0.08153**	**0.07231**

Table 4.5: HOMO-LUMO Egap (in hartrees) for the five polyyne molecules. A decreasing trend with increase in chain lengths is observed. The Egap also decreases with increasing bending angle for each molecule.

Conclusions

The studies conducted in this chapter explored the mechanical and electronic properties of polyynes, using both semi-empirical and DFT approaches. The computations dealt with the changes in these properties as a result of bending and stretching of these molecules. Thus, we demonstrated for the first time, the variations in total energy and electronic properties of bent polyyne molecules. In all cases, we found that the slopes of the total energy curves decrease with increasing chain lengths, which indicates that the force constants are different depending on chain lengths. We also found that the molecules are bent in such a way that the degree of bending is highest at the C_1-C_2-C_3 bonds and decreases further along the chain, and that the overall extent of bending decreases with increasing chain lengths, as revealed by the C-C-C- bond angles and the slopes of their curves. In addition, we discovered that for DFT simulations, the decrease in degree of bending along the chain is gradual while for PM3 simulations, this decrease is rather drastic. Thus, the bending is relatively spread out across the entire chain for DFT calculations, while for PM3 approach, bending is rather concentrated in the C_1-C_2-C_3 bonds and drastically decreases further along the chain. Hence, the molecules are not uniformly bent in this case.

For both PM3 and DFT calculations, it was found that the stretching force constant is about two orders of magnitude times the bending force constants, meaning that more force is needed to stretch the molecules than to bend them. We observed a similar relationship between bending and stretching force constants, which we determined from the experimentally observed wavenumbers of DP-C12. Thus, higher energy is required to cause stretching than bending, indicating the higher force constants for stretching than bending. We used the K_{eff} functions to fit the force constant curves and reasonably good fits were obtained. Thus, the molecules may be said to behave to a first approximation, as units of springs joined together. However, the values of the fitting parameters and force constants obtained from the fits are different from those originally estimated in chapter 3. The differences are more significant for bending curves than for stretching. This may be due to the fact that our assumption on the spring-like behaviour of the molecules is inadequate. This might also be due to inaccuracy of our bending calculations or poor assignments of the IR peaks. Since the discrepancy is small in the case of stretching, the approximation is more valid for the stretching calculations. However, the fact that k_2 and k_3 have essentially same value indicates that the molecules are effectively conjugated across the entire chain lengths.

The force constant curves were also fitted with flexural rigidity function which approximates the molecules as flexible rods. In all cases, better fits were obtained with this function than the Keff function. Thus, it may be argued that this latter approximation is more adequate for describing the mechanical behaviour of these molecules. For this function however, we noted significant discrepancies between our fitting parameters and their original constant values. This may be ascribed to the fact that our approximation of the molecules as flexible rods is less adequate, or it may be because the original constant values are themselves results of computations and mechanical engineering approximations, which in themselves may be inadequate.

The force constants obtained for these molecules are characteristic of their chain lengths. We also found that the bending force constants of some of the molecules fall in the range of some AFM cantilevers' force constants. Thus, we predict that the responses of these molecules under the influence of an AFM tip will depend on their relative force constants. In addition, it was observed that surface-bound molecule has contributions to the total energy from the silicon and hydrogen atoms representing the surface, this changes the energy and force constants from those observed for the free unbound molecule.

We also found that the Egap decreases with chain lengths and bending angle, but the decrease is independent of changes in equilibrium length of these molecules. When the normalized Egap was plotted as a function of the average atomic angle, we found no difference between the different chain lengths. Since DFT results are more accurate than PM3 method, the DFT calculations and results are deemed more accurate than PM3 calculations. The results from these calculations can serve as a springboard for further experimental exploration of the mechanical and electronic properties of these molecules.

References

1. E. Kloster-Jensen, H.-J. Haink and H. Christen, *Hel. Chim. Acta,* 57, 1731 (1974).
2. A. D. Slepkov, et al. *J. Chem. Phys.* **120,** 6807 (2004).
3. A. Arnau, I. Tunon, E. Silla and J. M. Andres, *J. Chem. Ed.* **67,** 905 (1990).
4. M. Tsuji, S. Kuboyama, T. Matsuzaki, and T. Tsuji, *Carbon,* **41,** 2141 (2003).
5. L. Horny, N. D. K. Petraco, C. Pak and H. F. Schaefer, *J. Am. Chem. Soc.* **124,** 5861 (2002).
6. M. Grutter, M. Wyss, J. Fulara and J. P. Maier, *J. Phys. Chem. A,* **102,** 9785 (1998).
7. A. Scemama, P. Chaquin, M.-C. Gazeau and Y. Benilan, *Chem. Phys. Lett.* **361,** 520 (2002).
8. J.P. Maier, *J. Phys. Chem. A,* **102,** 3462 (1998).
9. A. Scemama, P. Chaquin, M.-C. Gazeau and Y. Benilan, *J. Phys. Chem. A,* **106,** 3828 (2002).
10. F. Pauzat and Y. Ellinger, *Astron. Astrophys.* **216,** 305 (1989).
11. J. Cernichazo and M. Guelin, *Astron. Astrophys.* **309,** L27 (1996).
12. M. Guelin, et al. *Astron. Astrophys.* **317,** L1 (1996).
13. T. D. Poulsen, et al. *J. Chem. Phys.* **114,** 5917 (2001).
14. U. Molder, P. Burk and I. A. Koppel, *Int. J. Quant. Chem.* **82,** 73 (2001).
15. S. Eisler, et al. *J. Am. Chem. Soc.* **127,** 2666 (2005).
16. M. A. Heuft, S. K. Collins, G. P. A. Yap and A. G. Fallis, *Org. Lett.* **3,** 2883 (2001).
17. H. S. Nalwa, J. Mukai and A. Kakuta, *J. Phys. Chem.* **99,** 10766 (1995).
18. J. L. Toto,T. T. Toto and C. P. de Melo, *Chem. Phys. Lett.* **245,** 660 (1995).
19. C. K. Chiang, et al. *Phys. Rev. Lett.* **39,** 1098 (1977).
20. A. D. Slepkov, et al. *J. Chem. Phys.* **116,** 3834 (2002) .
21. F. Coat and C. Lapinte, *Organometallics,* **15,** 477 (1996).
22. D. W. Rogers, et al. *J. Org. Chem.* **69,** 7143 (2004).
23. T. Luu, et al. *Org. Lett.* **7,** 51 (2005).

24. S. Szafert and J. A. Gladysz, *Chem. Rev.* **103,** 4175 (2003).

25. Q. Fan and G. V. Pfeiffer, *Chem. Phys. Lett.* **162,** 472 (1989).

26. *HyperChem 7.5, a molecular modeling product from Hypercube, Inc.* , 1115 NW 4th St., Gainesville, FL 32601

27. M. J. Frisch, et al. *Gaussian 03, Revision B.05,* Gaussian, Inc., Pittsburgh PA, 2003.

28. *http://www.chem.ualberta.ca/~abrown/*

29. *http://www.westgrid.ca/support_old/software/gaussian.php*

30. M. J. Frisch, et al. *Gaussian 03, Revision C.02,* Gaussian, Inc., Pittsburgh PA, 2003.

31. An unpublished result (chapter 2 of this thesis) where the performance of fifteen different commonly used functionals on vibrational frequency of a model polyyne 1, 12-Diphenyl-1, 3, 5, 7, 9, 11-dodecahexayne were explored. BLYP has the overall best performance. In a related but separate work, bending calculations performed on different polyyne molecules with B3LYP and BLYP functionals gave similar total energies and exactly the same bending force constants for the polyynes, but BLYP functionals takes less computation time.

32. X. Zhou, C.J.M. Wheeless and R. Liu, *Vibrational Spectroscopy,* **12,** 53 (1996).

33. X. Zhou, S.J. Mole and R. Liu, *Vibrational Spectroscopy,* **12,** 73 (1996).

34. C. J. M. Wheeless, X. Zhou and R. Liu, *J. Chem. Phys.* **99,** 12488 (1995).

35. J. Higgins, X. Zhou and R. Liu, *Spectrochim Acta Part A,* **53**, 721 (1997).

36. X. Zhou and R. Liu, *Spectrochim Acta Part A,* **53,** 259(1997).

37. B. Ball, X. Zhou and R. Liu, *Spectrochim Acta Part A,* **52**, 1803 (1996).

38. B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.* **98,** 5612 (1993).

39. A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).

40. C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B,* **37**, 785 (1988).

41. *B. Miehlich, A. Savin, H. Stoll, and H. Preuss, Chem.* Phys. Lett. **157**, 200 (1989).

42. I. Navizet, F. Cailliez and R. Lavery, *Biophysical Journal,* **87,** 1426 (2004).

43. J. Howard, Mechanics of Motor Proteins and the Cytoskeleton, Sinauer Associates, Publishers, (2001).

44. L. Itzhaki, E. Altus, H. Basch and S. Hoz, *Angew. Chem. Int. Ed.* **44,** 7432, (2005).

Chapter 5

General Conclusions

This research work was aimed at spectroscopic characterization of oligonucleotides immobilized on gold on the one hand, and on the other hand, probing the mechanical and electronic properties of single molecules i.e. polyynes. These form the subjects of discuss in chapters 2 to 4.

The work presented in Chapter 2 demonstrates the immobilization of ssDNA on gold surface, as well as their characterization with XPS and FTIR [1, 2]. It must however be noted that other methods are being used by different researchers for characterizing these molecules [3-7]. If proper film parameters are employed and realistic assumptions are made, XPS can be devised as a tool for quantitative and qualitative characterization of the surface-bound ssDNA, as we demonstrated in Chapter 2. In the same vein, FTIR can give qualitative information about the oligonucleotides on the gold surface. Thus, both techniques were used in this work as complementary techniques for ssDNA characterization.

The results from our XPS characterization give some evidence on the adsorption of the ssDNA molecules on the surface. Based on the S2p signal, our XPS results suggest that either one or both of these two events occurred for sample TT: the sample was adsorbed onto the gold surface via the thiol group which is buried under the overlying DNA molecules, hence the very weak signal observed; the sample coadsorbed with certain contaminants which sufficiently buries the sulphur atom and makes the surface inaccessible to MCH, resulting in the very weak S2p signal. As for sample AA, the molecules were believed to be chemisorbed on the gold surface non-specifically, and when backfilled with MCH, some AA molecules were displaced off the surface such that the S2p signal remained essentially the same. This observation indicates that MCH being a short chain alcohol can be effective at controlling the surface density of immobilized ssDNA and in removing non-specific adsorption [5]. Quantitative analysis was carried out using the N1s and Au4f signals and the overlayer formalism of Petrovykh [1] to determine the film thickness and relative atomic densities for samples TT and AA. TT was found to be a relatively thicker film than AA, which may be due to coadsorption of contaminants with TT on the surface, while AA has a higher relative atomic density due to its higher N1s signal. The surface coverage could not be reliably determined because of the ambiguities observed in the data. Thus, more experiments have to be carried out on similar and replicate samples in order to obtain more reliable results, for only one sample was analyzed in each case here.

FTIR analysis also gives qualitative support for the adsorption of the ssDNA molecules. The peaks characteristic of carbonyls in free and chemisorbed thymine rings were found for sample TT, while for sample AA, we found peaks that correspond to NH_2 scissoring and C-N in-plane vibration, which are typical of chemisorbed adenine [1-2, 11]. Peaks corresponding to asymmetric and symmetric PO_2^- stretches were also found for both samples [2], further showing that some ssDNA molecules were adsorbed on the gold surface. It must however be mentioned that the results obtained from this work have to be further validated by carrying out further experiments on replicate samples to get a more definitive results which will corroborate or nullify the conclusions drawn from the work reported here. This is important, given the fact that.the results presented here are just from a single sample in each case.

Rational selection of a suitable DFT functional that can be used for modeling the mechanical and electronic properties of polyynes was the focus of the work presented in chapter 3. We used fifteen different DFT functionals (**SVWN, SPL, BP86, BLYP, BPW91, B3LYP, B3P86, B3PW91, B1B95, B1LYP, MPW1PW91, PBE1PBE, B98, B971 and B972**), and tested their performance at predicting the vibrational frequencies of our chosen model molecule 1, 12-Diphenyl-1, 3, 5, 7, 9, 11-dodecahexayne (DP-C12). We used a large basis set for most of these calculations since this gives more accurate results than the usual "spectroscopic" basis set (6-31G*). We used vibrational frequencies of ethylene as a benchmark, because this molecule has been extensively studied by various authors [12-13]. Our results from calculations on ethylene using BLYP, B3LYP and BPW91 in combination with 6-31G* basis set are in agreement with reported values (with an R^2

of 1.000). With the 6-31G** basis set, we got values that are even closer to experimental values than with 6-31G* basis set. In both cases, BLYP emerged the functionals of best overall performance [12-14], by giving values that are closest to experimental values. In this work, performance was used to describe the accuracy of the calculated values with respect to experimental values. Having established that our results support what other researchers have found for ethylene, we then used the fifteen functionals stated above on the chosen model molecule. Out of all the functionals tested with the chosen basis set, BLYP was found to have the best overall performance, with a percentage mean absolute deviation (MPAD) of 0.338%. One interesting finding from these calculations is that even with 6-31G* basis set, this functional with MPAD of 0.431%, also outperformed all other functionals used with the large basis set 6-31G** signifying the excellent performance of this functionals.

Generally, for both calculations on ethylene and DP-C12, we found that the performance of the functionals has correlation with the way they were formulated. The best functionals are the gradient-corrected functionals, followed by the local functionals, while the hybrid functionals performed the least. The gradient-corrected functionals - **BLYP, BPW91** and **BP86** performed best because they are based on the GGA approximation which gives correct qualitative description of molecular systems and correct the inhomogeneity inherent in the LDA approximation. Thus, functionals that are based on this approximation would be able to compute accurate vibrational frequencies. The local functionals **SVWN** and **SPL** gave less accurate frequencies since they are based on the rather less accurate LDA approximation. However, the hybrid functional **B98** is an exception in the list because its performance caused it to be grouped with the local functionals. On inspecting the way this functionals was formulated, we found that it was formulated in an analogous way to the local functionals – its exchange and correlation parts are composed of 75% or 100% local terms making it very similar to local functionals. Thus, this functional falls in the same performance rating as the originally formulated local functionals. The other hybrid functionals **B3LYP, B3P86, B3PW91, B1B95, B1LYP, MPW1PW91, PBE1PBE, B971 and B972** performed the least. We rationalized this level of performance on the basis of the fact that their exchange components contain HF, LDA and GGA terms, while their correlation parts consist of substantial fraction of local terms, which are rather less accurate. Due to the fact that the HF method and the LDA are both less accurate for qualitatively describing some molecular systems, which may include DP-C12, then the computed vibrational frequencies for our model molecular system may not be as accurate. Hence, the poor performance of these functionals for this molecule.

In terms of CPU time per iteration, the local functionals took about 80 minutes, while the gradient-corrected and most hybrid functionals (excluding those used with 6-31G*) took about 2 hours or more. This may be because the local functionals use the LDA approximation, which is not so accurate in describing molecular systems, and so took the least time at the expense of accuracy. The approximation used by the gradient-corrected functionals is more accurate and so these functionals require more time in finding the optimized geometry for the molecule and computing its vibrational frequencies. The fact that the hybrid functionals use blends of both approximations in conjunction with HF theory might have caused them also to take so long. Those functionals used with 6-31G* basis set took the least time (about 50 minutes of CPU time per iteration). This is ascribed to the lower level of theory used in this case.

From the experimentally observed wavenumbers on DP-C12, we obtained force constants and found the bending force constant to be about two orders of magnitude less than that of stretching. This same relationship was obtained from our computations on C_4-C_{12}, i.e. bending force constants were about two orders of magnitude less than stretching force constants. This is indicative of the relative energies required to cause these types of vibrations.

This work reports for the first time the performance of different functionals at predicting the vibrational frequencies of a polyyne molecule. Our work also confirms and extends the recommendations of various authors regarding the excellent performance of BLYP functionals. With these outcomes, we then used the BLYP functionals in combination with 6-31G** for modeling the mechanical and electronic properties of polyynes.

The results presented in chapter 4 are those of our exploration of the mechanical and electronic properties of polyynes – a class of compounds with unique properties [15-16]. We used a combination of PM3 and DFT approaches to model these properties. The kind of computations done here adds another dimension to the already existing data on straight-chain polyynes, for, we demonstrate for the first time, how the total energy of these compounds changes with bending and stretching. In addition, we examined how the bending might impact on their electronic properties by way of the HOMO-LUMO energy gap (Egap). We simulated the bending and stretching of members of straight chain polyynes and examined how these perturbations change their total energies. The PM3 calculations give rough estimates of the variation of the energy with bending angles and stretching coordinates, while DFT gives more accurate picture of these variations. For the straight unbent molecules, our DFT calculations give total energy results that are in excellent agreement with reported values to within 0.04% [17-18].

In all instances, we found that the slopes of the total energy curves decrease with increasing chain lengths, indicating that the force constants are chain length dependent. We also found that the molecules are bent such that the extent of bending is highest at the C_1-C_2-C_3 bonds and decreases further along the chain, and that the overall extent of bending decreases with increasing chain lengths, as revealed by the C-C-C- bond angles and the slopes of their curves. Also, we discovered that for DFT simulations, the decrease in degree of bending along the chain is gradual while for PM3 simulations, this decrease is more drastic. Therefore, the bending is relatively spread out across the entire chain for DFT calculations, while for PM3 approach, bending is rather concentrated in the C_1-C_2-C_3 bonds and drastically decreases further along the chain. Hence, the molecules are not uniformly bent in that case.

We also discovered that for both methods, the stretching force constant is about two orders of magnitude times the bending force constants, meaning that more force is needed to stretch these molecules than to bend them. An analogy can be drawn between this observation and the vibrational modes of molecules. They are similar cases because stretching modes occur at a higher wavenumber and require more energy than bending modes. Thus, higher energy is required to cause stretching than bending, indicating the higher force constants for stretching than bending. We obtained force constants for these molecules, which are characteristic of their chain lengths. We also found that the bending force constants of some of the molecules fall in the range of some AFM cantilevers' force constants. In addition, we estimated the effective bending and stretching force constants from the experimentally observed wavenumbers of DP-C12 for a set of polyyne molecules. We also derived useful functions based on the assumption that these molecules behave like springs. These functions were used to fit the force constant curves obtained from our computations. Good fits were obtained, showing that there may be some validity in our assumptions. However, the fitting parameters and force constants resulting from the fitting are different from the constant values estimated originally. This may mean that our assumption on the bending and stretching behaviour of the molecules is inadequate, though the fitting functions make some physical sense. It may also be concluded that this disagreement is due to inaccuracy of our bending calculations or poor assignments of the vibrational bending modes. Our approximation and assumption is however more valid for stretching than for bending, since the force constants obtained from the fitting are not significantly different from the original values.

We also fitted the force constant curves with flexural rigidity function [19], which approximates the molecules as flexible rods and is based on the flexural rigidity of the rods. In all cases, better fits were obtained with this function than the Keff function. We therefore suggest that this latter approximation is more adequate for describing the mechanical behaviour of these molecules. For this function however, we noted significant discrepancies between our fitting parameters and their original constant values. This may be ascribed to the fact that our approximation of the molecules as flexible rods is less adequate, or it may be because the original constant values are themselves results of computations and mechanical engineering approximations [20], which in themselves may be inadequate.

By manipulating the energy curves resulting from the PM3 stretching simulations, certain interesting trends in the intrinsic mechanical properties of these molecules were revealed. This became apparent when we obtained a plot of stretching force vs

% stretched length by normalizing the stretching coordinates of each molecule to its unstretched length. The plot enabled us to see the intrinsic differences between the members of polyyne series considered, apart from chain length differences. This manipulation enabled us to see the variation of the stretching force as a function of strain for each molecule. This manipulation was not carried out on the DFT energy curves as only five members were considered and no definite pattern can be observed for such a small group of molecules, compared with the nineteen molecules treated in the case of PM3 stretching.

We also explored the electronic properties of these molecules and found the HOMO-LUMO Egap to decrease with chain lengths. In addition to this, the normalized Egap when plotted as a function of average atomic angle, was found to be independent of chain length.

The calculations performed on the more realistic surface-bound polyyne indicate that the silicon and hydrogen atoms representing the surface contribute to the total energy, this changes the energy and force constants from those observed for the free unbound molecule. The electronic properties of both are also likely to be slightly different. This is based on the fact that the HOMO-LUMO Egap of the surface bound molecule is lower than that of the free molecule by a constant value of about 0.005 hartrees for each bending angle. Thus, under the influence of an AFM tip, these molecules are expected to bend in such a way that their electronic properties are also influenced.

These results, we believe, could serve as a frame of reference for further experimental work to practically examine the properties we have considered here. Then it can be seen if the deductions made here are valid or otherwise.

References

1. D. Y. Petrovykh, H. Kimura-Suda, M. J. Tarlov and L. J. Whitman, *Langmuir*, **20,** 429 (2004).

2. D. Y. Petrovykh, H. Kimura-Suda, L. J. Whitman and M. J. Tarlov, *J. Am. Chem. Soc.* **125,** 5219 (2003).
3. A. W. Peterson, R. J. Heaton and R. M. Georgiadis, *Nucl. Acid. Res.* **29,** 5163 (2001).
4. A. B. Steel, R. L. Levicky, T. M. Herne and M. J. Tarlov, *Biophys. J.* 79, 975 (2000).
5. T. M. Herne and M. J. Tarlov, *J. Am. Chem. Soc.* **119,** 8916 (1997).
6. E. A. Smith, et al. *Langmuir,* **17,** 2502 (2001).
7. C. Y. Lee, H. E. Canavan, L. J. Gamble and D. G. Castner, *Langmuir,* **21,** 5134 (2005).
8. L. K. Wolf, Y. Gao, and R. M. Georgiadis, *Langmuir,* **20,** 3357 (2004).
9. H. Kimura-Suda, D. Y. Petrovykh, M. J. Tarlov and L. J. Whitman, *J. Am. Chem. Soc.* **125,** 9014 (2003).
10. L. M. Demers, et al. *J. Am. Chem. Soc.* **124,** 11248 (2002).
11. D. Y. Petrovykh, et al. *J. Am. Chem. Soc.* **128,** 2 (2006).
12. X. Zhou, C. J. M. Wheeless and R. Liu, *Vibrational Spectroscopy,* **12,** 53 (1996).
13. B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.* **98,** 5612 (1993).
14. C. J. M. Wheeless, X. Zhou and R. Liu, *J. Chem. Phys.* **99,** 12488 (1995).
15. A. D. Slepkov, et al. *J. Chem. Phys.* **120,** 6807 (2004).
16. T. D. Poulsen, et al. *J. Chem. Phys.* **114,** 5917 (2001).
17. U. Molder, P. Burk and I. A. Koppel, *Int. J. Quant. Chem.* **82,** 73 (2001).
18. L. Horny, N. D. K. Petraco, C. Pak and H. F. Schaefer, *J. Am. Chem. Soc.* **124,** 5861 (2002).
19. J. Howard, Mechanics of Motor Proteins and the Cytoskeleton, Sinauer Associates, Publishers, (2001).
20. L. Itzhaki, E. Altus, H. Basch and S. Hoz, *Angew. Chem. Int. Ed.* **44,** 7432, (2005).

Appendix

Computation Methods

Hartree-Fock (HF) Method: This is a method for computing various molecular properties based on solutions to the Schröndinger equation. Because this method attempts to solve the equation exactly, it is termed *ab initio* method. This method is an all-electron method and gives good descriptions for most molecular systems. Thus, the results are believed to be accurate, but it takes a lot of CPU time.

Semiempirical Methods: These are based on approximations to the *ab initio* HF method. The approximations involved the use of experimental parameters and the neglect of most 1 and 2-electron integrals that are normally computed in HF calculations. In addition, they use only valence electrons and minimal basis set in their computations. Thus, these methods are used to perform computations on large molecules in reasonable CPU time. Examples include AM1 and PM3 methods.

DFT Methods: These are also based on some approximations to the exact HF method. In addition, these methods include electron correlation in their computations. This is an added advantage for these methods, as the HF method lacks electron correlation. These methods are based on the assumption that ground state molecular properties can be determined from the ground state electron density. This means that if the ground state electron density of a molecular system can be determined, then all ground state molecular properties can be obtained from this electron density. The electron density is expressed by exchange-correlation functionals, from which the ground state molecular properties are evaluated.

Similarities and Differences

All the three methods can be used to perform computations on small molecular systems, e.g. hydrogen and water molecules. They are reasonably accurate at predicting bond lengths. However, some are better for computing certain other molecular properties than others. For instance, DFT methods give better results on properties that involve electron density and electron correlation effects, while semiempirical methods are not good for properties that depend on nuclear properties e.g. NMR shielding parameters.

HF method is accurate because it is based on finding solutions to the Schröndinger equation but it is computationally expensive, especially for large systems. However, this method does not include electron correlation effects. Semiempirical methods are good for large systems but may give less accurate results for many properties. However, since their parameters are derived from experiments, they include electron correlation. DFT methods can achieve accuracy close to or even better than HF method, depending on which functionals are used. With a very good choice of functionals, these methods can give accurate results in less time compared to HF method. But, since the assumptions are for ground state systems, DFT methods are unsuitable for evaluating excited state properties.

DFT Functionals used for Calculations

As stated in the text, DFT methods make use of exchange-correlation functionals in their calculations. The DFT expressions for the electron density are composed of two terms: exchange term and correlation term. These are based on certain approximations, viz: local density approximation (LDA) and gradient corrected or generalized-gradient approximation (GGA). The LDA assumes a slowly varying electron density as a function of position, while GGA includes gradient of the density in its approximation. Another type of functionals is hybrid functionals, which mixes together HF, local and gradient terms.

Local Functionals

These have both exchange and correlation parts made of local terms.

SVWN: Combines Slater local exchange with local correlation functional of Perdew.

SPL: Slater exchange functional plus Vosko, Wilk and Nusair 1980 correlation functional.

Gradient-corrected Functionals

Both exchange and correlation parts made of gradient-corrected terms.

BP86: Becke's gradient-corrected exchange plus Perdew's correlation functional.

BLYP: Becke's exchange plus correlation functional developed by Lee, Yang and Parr.

BPW91: Becke's exchange combined with correlation functional of Perdew and Wang.

Hybrid Functionals

These are composed of HF-derived term, local and gradient-corrected terms. These terms have fitting parameters, which are constants derived by fitting experimental results with calculations. Their acronyms usually have numbers which indicate the number of fitting parameters used to derive the functionals.

B3LYP: Becke's three-parameter hybrid functional combining Becke's exchange with Lee, Yang and Parr's correlation terms.

B3P86: A three-parameter hybrid functional combining Becke's exchange with Perdew's correlation.

B3PW91: Three-parameter functional of Becke's exchange plus correlation functional of Perdew and Yang.

B1B95: A one-parameter hybrid functional consisting of Becke's exchange and correlation.

B1LYP: A one-parameter functional combining Becke's exchange with the correlation functional of Lee, Yang and Parr.

MPW1PW91: A one-parameter hybrid functional which modifies Perdew-Wang's exchange and correlation.

PBE1PBE: One-parameter hybrid functional formulated by Perdew, Burke and Ernzerhoff.

B98: A hybrid functional of Becke, which is a revision to his B97 functional.

B971: Handy, Tozer and coworkers' modification to Becke's B97 functional.

B972: Wilson, Bradley and Tozer's modification to B97.

Computed Wavenumbers

Out of the fifteen functionals used for computing the wavenumbers of our model molecule DP-C12, four functionals were used with 6-31G* basis set to ensure convergence to a local minimum. When these functionals, were used with 6-31G** basis set, the computations converged to a saddle point, as indicated by negative wavenumbers. Below are listed all the wavenumbers in units of cm^{-1} computed by these functionals at the 6-31G** level. Each of the functionals gave one negative wavenumber.

B3LYP

-1.8871	10.0005	10.8871
28.2242	32.3897	53.9984
63.6314	92.3261	103.6484
148.1456	153.1773	160.9368
211.4626	220.6632	270.8136
296.3274	316.3535	351.6238
382.7537	409.3800	409.5895
456.3129	477.6159	486.5083
495.8756	505.7416	526.4029
564.3284	575.2047	634.0270
635.0613	640.2790	655.6849
701.6992	701.8065	707.2513
732.1133	768.5262	773.8182
801.6212	834.1351	855.5460
855.5472	899.3273	933.9301
934.5302	976.3605	976.3608
996.6080	998.3327	1003.5511
1003.7938	1013.8724	1014.7898
1052.3377	1055.1675	1081.5623
1107.9984	1108.0021	1137.0693
1159.8091	1189.7557	1189.7558
1205.9985	1207.4041	1279.6140
1324.7869	1324.7876	1361.4202
1361.4203	1392.1014	1481.5760
1481.5771	1508.6870	1535.2527
1554.6163	1619.5773	1619.5778
1650.7660	1651.0109	2127.9607
2128.7645	2212.4428	2260.0857
2276.2444	2299.4674	3185.0644
3185.0792	3194.0425	3194.0428

3204.3555	3204.3670	3213.0663
3213.0688	3216.4052	3216.4612

B3P86

-4.5914	10.5395	10.7916
29.0080	32.2007	55.5846
63.3701	94.4768	103.2097
150.0023	152.3700	161.9555
210.1842	221.6530	269.2610
295.5350	317.5359	351.1782
381.7883	406.7642	406.9855
453.5929	480.4897	487.0879
491.6724	514.0707	525.4278
564.0469	584.1723	631.4506
633.1325	640.0961	653.8513
701.3129	701.4542	718.3612
733.8847	766.5704	772.8631
803.0875	839.0883	854.7431
854.7445	917.8406	934.0231
934.7633	978.4725	978.4727
1003.6308	1003.6902	1004.5577
1006.5089	1015.3380	1015.5241
1058.2191	1060.9986	1103.6157
1111.4056	1111.4061	1151.2615
1167.2638	1187.8685	1187.8690
1204.5022	1206.4078	1290.2270
1332.7395	1332.7399	1376.6978
1376.6978	1403.5468	1484.5426

1484.5444	1513.8469	1539.4766
1567.4686	1634.1327	1634.1333
1666.2355	1666.5558	2135.6164
2144.4965	2223.1265	2276.2549
2289.3039	2314.4261	3202.0512
3202.0656	3211.0938	3211.0941
3220.9839	3220.9920	3229.4081
3229.4098	3232.7799	3232.8293

BP86

-3.5922	10.0890	10.5522
26.7234	31.2710	49.6646
61.1979	85.5249	99.5120
139.6577	147.2122	158.9064
203.3214	209.9406	260.7028
284.2313	309.9140	333.1595
364.3331	391.0445	391.4439
432.9693	448.6270	463.4379
467.0720	475.2405	501.4101
521.7308	529.0888	598.2875
610.7029	611.8559	623.2366
626.6352	674.5188	674.5896
706.3063	713.3005	740.6619
743.4315	785.0242	817.5162
817.5691	817.5778	885.7135
890.9163	896.1333	917.4258
931.4692	931.4731	957.6553
957.6986	978.1076	981.0838
982.3279	984.3220	1022.7794

1025.6515	1075.9602	1075.9607
1136.3199	1155.0666	1155.0671
1169.7086	1172.2939	1259.4293
1289.0661	1289.0667	1341.0705
1341.0705	1380.9105	1433.4956
1433.4968	1468.4386	1485.5552
1554.1234	1563.2979	1563.2985
1595.8092	1596.7682	2020.6343
2024.3793	2123.9520	2171.8203
2193.6176	2216.4475	3106.8663
3106.8878	3115.4357	3115.4361
3126.4616	3126.4721	3135.1911
3135.1938	3138.3429	3138.4247

B972

-3.1809	10.7991	11.2463
29.9243	32.0496	57.3113
63.0475	96.9276	102.6621
151.3803	152.0986	161.7818
208.3979	221.9893	266.7098
292.6033	317.5408	348.8665
377.2164	407.1013	407.3035
449.3625	483.9840	485.1647
487.4080	519.8634	523.3779
555.2749	580.8098	629.9237
632.5885	634.5402	640.6832
702.6571	702.7768	704.0038
734.0908	753.8769	773.7030
780.3331	838.5845	858.2315

858.2323	889.0453	935.7499
938.2076	950.6769	983.4020
983.4025	1002.7320	1008.5208
1008.5428	1016.4554	1016.5613
1060.5175	1063.7833	1069.8732
1098.9615	1115.6446	1115.6450
1167.2457	1193.6768	1193.6773
1211.0064	1212.4919	1289.0489
1337.3756	1337.3774	1376.8286
1376.8288	1401.4334	1489.0323
1489.0342	1515.5721	1543.6924
1565.5081	1638.0477	1638.0483
1670.0336	1670.2930	2140.8677
2148.9244	2226.6283	2278.6783
2291.9012	2316.6797	3216.8908
3216.9051	3226.2125	3226.2128
3236.4478	3236.4572	3245.1316
3245.1338	3248.5521	3248.6086